



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 117364

To: **Phuong Bui**
Location: **rem 2a15 & 2c18**
Art Unit: **1638**
Tuesday, March 23, 2004

Case Serial Number: **09/913064**

From: **Beverly Shears**
Location: **Remsen Bldg.**
RM 1A54
Phone: **571-272-2528**

beverly.shears@uspto.gov

Search Notes

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 17:29:48 ; Search time 5324 Seconds
(without alignments)
10681.078 Million cell updates/sec

Title: US-09-913-064a-13
Perfect score: 1312
Sequence: 1 Gcagcagcactctctctcc.....aaaaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 347272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.vrt.*
- 37: em.htg.man.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	833.2	63.5	1216	8	PSU31544	U31544 Pisum sativ
2	673.2	51.3	1474	8	CTAU5081	AJ005081 Cyamopsi
3	649.2	49.5	1170	8	BT000032	BT000032 Arabidops
4	649.2	49.5	1313	8	AY054207	AY054207 Arabidops
5	649.2	49.5	1420	8	AY120709	AY120709 Arabidops
6	644.4	49.1	1056	8	AX507662	AX507662 Sequence
7	644.4	49.1	1056	8	AF334724	AF334724 Arabidops
8	642.8	49.0	1356	8	ATUDGE	Z54244 A.thaliana
9	631	48.1	1263	8	AK117913	AK117913 Arabidops
10	629.4	48.0	1397	8	AY085887	AY085887 Arabidops
11	621.6	47.4	1240	8	AY094481	AY094481 Arabidops
12	482.8	36.8	1375	8	AK104977	AK104977 Oryza sat
13	482.8	36.8	1425	8	AK073610	AK073610 Oryza sat
14	479.6	36.6	1223	8	AB097460	AB097460 Oryza sat
15	476.4	36.3	1516	8	BT009044	BT009044 Triticum
16	456.4	34.8	1360	8	CTAU5082	AJ005082 Cyamopsi
17	445	33.9	1065	6	AX653074	AX653074 Sequence
18	445	33.9	1346	8	AB087745	AB087745 Oryza sat
19	435	33.2	1736	8	AB096863	AB096863 Oryza sat
20	435	33.2	1878	8	AK066269	AK066269 Oryza sat
21	433.4	33.0	1420	8	AF303682	AF303682 Zea mays
22	428.4	32.7	1286	8	AY085528	AY085528 Arabidops
23	423.6	32.3	1087	8	AY117180	AY117180 Arabidops
24	423.6	32.3	1356	8	AY065354	AY065354 Arabidops
25	423.6	32.3	2800	8	AY140073	AY140073 Arabidops
26	416.8	31.8	1316	8	AY084615	AY084615 Arabidops
27	416.4	31.7	1419	8	AK121426	AK121426 Oryza sat
28	416	31.7	1084	8	BT008539	BT008539 Arabidops
29	416	31.7	1393	8	AK118722	AK118722 Arabidops
30	385.6	29.4	1335	6	AX653761	AX653761 Sequence
31	313.6	23.9	10398	1	AE015612	AE015612 Shewanell
32	296.2	22.6	276034	1	AE017015	AE017015 Bacillus
33	287	21.9	17317	1	AY147913	AY147913 Streptoco
34	278.2	21.2	4626	14	AF065660	AF065660 Stealth v
35	277.8	21.2	1017	6	BD170587	BD170587 UDP-N-Ace
36	277.8	21.2	6715	1	BSGALE	X93339 B.subtilis
37	277.8	21.2	65143	1	D83026	D83026 Bacillus su
38	277.8	21.2	201139	1	BSUB0020	Z99123 Bacillus su
39	271.6	20.7	306050	1	BX321858	BX321858 Nitrosomo
40	270.6	20.6	250150	1	AP005342	AP005342 Vibrio vu
41	269.8	20.6	11351	1	AE004405	AE004405 Vibrio ch
42	267.6	20.4	1390	3	AK114505	AK114505 Ciona int
43	267.6	20.4	69727	6	AX770896	AX770896 Sequence
44	267.6	20.4	124056	1	BX571875	BX571875 Photornab
45	267	20.4	302000	1	AP003187	AP003187 Clostridi

ALIGNMENTS

RESULT 1	PSU31544	1216 bp	mRNA	linear	PLN 02-FEB-1996
LOCUS	Pisum sativum	UDP-galactose-4-epimerase (gale)	mRNA	complete cds.	
DEFINITION	U31544				
ACCESSION	U31544				
VERSION	U31544.1	GI:1173554			
KEYWORDS					
SOURCE	Pisum sativum (pea)				
ORGANISM	Pisum sativum				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;				
	Pisum.				
REFERENCE	1 (bases 1 to 1216)				
AUTHORS	Slocum,R.D., Lake,M.R. and Williamson,C.L.				

TITLE Cloning and characterization of a UDP-galactose-4-epimerase ('galactowaldenase') and its expression in pea tissues

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1216)

AUTHORS Slocum, R.D.

TITLE Direct Submission

JOURNAL Submitted (12-JUN-1995) Robert D. Slocum, Biological Sciences, Goucher College, Baltimore, MD 21204-2794, USA

FEATURES

Location/Qualifiers

1..1216

/organism="Pisum sativum"

/mol_type="mRNA"

/cultivar="wandoPea"

/db_xref="taxon:3888"

1..1216

/gene="gale"

6..1058

/gene="gale"

/EC number="5.1.3.2"

/function="catabolism of galactose to glucose in Leloir pathway, and in galactose synthesis from glucose."

/note="galactowaldenase"

/codon_start=1

/product="UDP-galactose-4-epimerase"

/protein_id="AAA86532.1"

/db_xref="GI:117355"

/translation="WVASSOKILVTGSAGFIGTHVTWLLNNGFNVSIIIDFNSVME AVERREVGVNSLONLPTGLDKNDLEKLSKEDAVIHPAGLKVAGSVENP RYFDNLVGTNLNLEVNAKCNKWFSSSNATYGOPEKIPCTVEDFKLOANNPYGT KLFLBEADIDQKAPSPWRIVLIRYFNPVGAHESKLGEDPRGFINNMFLIQVAVG RLPELVYGHDPTRDGSAIRDYIHVMDLADGHIARLKLFTSENIENCTATNLGTRG SSVLEWAAFEKASGKKALKLCPRRPGDTEVYASTATAKELGKAKYGVEMCRD QWNWAKNNPWGYSKGP"

ORIGIN

Query Match 63.5%; Score 833.2; DB 8; Length 1216;

Best Local Similarity 84.0%; Pred. No. 1.7e-191;

Matches 953; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

QY 31 AGCATGGTGTCTTCCTCCCAACACATTCGTGGTCACCGGTGGTGGCGGTTTCATTGGCAC 90

DB 3 AACATGGTGGCTTCGTGCGAAGATACCTGTAACCGGCGAGTGGCGGTTTCATCGGCACT 62

QY 91 CACACGGTGTTCAGCTTCTCAAGCTGGCTTCAGGTTTCAATATCGACATTTCCGAT 150

DB 63 CACACGGTGTTCAGCTTCTCAATATCGGCTTCACGTTTCCATCATTAATTAATTCGAT 122

QY 151 AACTCCGTATGAAGCAGTGGACCGGTCGCGCAAGTGGTGGCCCTCTGCTTTCTCAG 210

DB 123 AATTCTGTATGAAGCAGTGGACCGGTCGCGCAAGTGGTGGCTTCAACCTCTCTCAG 182

QY 211 AACTCCAAATCCACCGGCGATCTCCGGAATAGGATGACTTGGAGAACTCTCTCTCC 270

DB 183 AATCTCGAATTCATCTGGGAGATCTCAGGAATAAAGATGATTTGGAGAACTCTCTCTCA 242

QY 271 AAAACACATTTGATGCGGTATCCATTTGCTGGCTTGAAGCGGTTCCTGAAGCGTT 330

DB 243 AATCTAAATTTGATGCTGTATTCATCTTGTGGAATAAGCAGTGGTGGAGTGT 302

QY 331 GCGAAGCCCGTGTCTATTTTGAATTTTAAATTTGGTGGCAACATCACTCTACGAGTTT 390

DB 303 GAAATCTCTGTCTTATTTTCGATAATACTTTGTTGGAATATCAATCTCTATGAAGTT 362

QY 391 ATGCAAGATTAATTCGAAGAGTGGTTTCTCATCTCGCAACCGTTTATGGCCAA 450

DB 363 ATGGCTTAAGCATTAATTTGAAGAGTGGTTTCTCGTCACTCGCAACTGTTTATGCCAA 422

QY 451 CCTGAAAGATACCGTGTGAGGAGATTTCAAGTTACAGCTATGAATCCCTATGACCG 510

DB 423 CCTGAAAGATACCGTGTGAGGAGATTTCAAGTTACAGCTATGAATCCCTATGACCG 482

QY 511 ACCAAGCTTTCTCGAAGAAATTTGCCCGGAGATATTCAGAAAGCTGAACAGATGAAG 570

DB 483 ACCAAGCTTTCTCTTGAAGAAATTCACGAGATATTCAGAAAGCTGAGCCAGATGGAGA 542

QY 571 ATCATATTACTAGATATCTCAATCAGTTGGGCTCATGAAGTGGCAAACTCGGTGAA 630

DB 543 ATCGTTTACTGCGGTACTTCAATCAGTCGCGGACATGAAGTGGTAACTTGTGTAA 602

QY 631 GATCCCAAGGGCATCCCAATAACCTTCATGCTTACATTCAGCAAGTACTGTGTGAAGA 690

DB 603 GATCCCAAGGGCATCCCAATAATCTCATGCTTATATACAGCAAGTACCGCTGGAAGA 662

QY 691 TTGACTGAATCAATATATACGGTCAATGATTATCCAAACAGGATGGCTTCGGATCCGG 750

DB 663 TTACCTGAGCTCAATGATATATGCTCATGATTATCTTCAAGGATGGCTTCGCGATACGG 722

QY 751 GACTATATCCATGTGATGACTTGGCAGATGGCCATATTCCTGCTCCCTGCGAAGACTCTTC 810

DB 723 GACTATATCCATGTGATGACTTGGCAGATGGCTTACATGCTGCTGCGCTGAGAAAGCTTTTC 782

QY 811 ACAACGGAGAACATAGGTTGATCTTACACCTGGGAACCTGCTGCTGCGAAGACTCTTC 870

DB 783 ACATCAGAAACATTCGTTGATCTGCTTATATTTGGGAACCTGCTGCTGCTCATCTGTG 842

QY 871 CTTGAAATGGTTACAGATTTGAAAGGCTTCGGCAAGAAATTCAGTAAATATATGT 930

DB 843 CTTGAAATGGTTGCTGCAATTTGAAAGCTTCGGCAGAAATTCGATGAAATTTGT 902

QY 931 CCAAGAGAACCGGAGATCGACTGAGTTTATGATCTTACAGAGAGACTGAGAAAGAA 990

DB 903 CCAAGAGAACCGGAGATCGACTGAGTTTATGATCTTACAGAGAGACTGAGAAAGAA 962

QY 991 CTTGTTTGGAGGCAACATTCGTTGGAGGATGTCGAGGACCAATCGAATTTGGGCA 1050

DB 963 CTTGTTTGGAGGCAACATTCGTTGGAGGATGTCGAGGACCAATCGAATTTGGGCA 1022

QY 1051 AAGAACAAATCCCTGGGTTACGGGGAGCCCTTGAATTAAGTTG---AGAAATATACTG 1107

DB 1023 AAGAACAAATCCCTGGGTTTACTCAGGAAGCTTGAATCAGTTGACATAAATACTA 1082

QY 1108 CTGATCTACGATCTTTTACATAATAGGCTCTCTTATATAGATACTTTT 1161

DB 1083 CACATATGCCAATCTTTCTGTACAAATAGGCTCCCATTTGTTAAGATACTT 1136

RESULT 2

CTAJ5081 1474 bp mRNA linear PLN 27-APR-1999

LOCUS Cyamopsis tetragonoloba mRNA for UDP-galactose 4-epimerase, clone

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..1474

/organism="Cyamopsis tetragonoloba"

/mol_type="mRNA"

/db_xref="taxon:3832"

1 Brunstedt, J., Joerbo, M., Pedersen, S.G. and Marcussen, J.

Isolation and expression of two cDNA clones encoding UDP-galactose

epimerase expressed in developing seeds of the endospermous legume

guar

Plant Sci. 142, 147-154 (1999)

2 (bases 1 to 1474)

Brunstedt, J.

Direct Submission

Submitted (24-MAR-1998) Danisco Biotechnology, Danisco A/S,

Langebrogade 1, PO Box 17, Copenhagen DK-1001, Denmark

Location/Qualifiers

```

CDS
/clone="GEPI42"
79..1143
/BC_number="5.1.3.2"
/codon_start=1
/product="UDP-galactose 4-epimerase"
/protein_id="CAA06338.1"
/db_xref="GI:3021355"
/db_xref="GOA:O65780"
/db_xref="SWISS-PROT:O65780"
/translation="WYSSRKASGEHILVGGAGFVSHTVVQLLKGGHVSILDLNLYN
SVLDVHVRLLVGLPDLNHLNLDLILFSLKTFDAVHIFAGLKGVGES
VLNPSNYDNVILNLFQVMSKFNKLLVSSATVGGPDQIPCVEDSNLHAWNP
YGSKLFVEVARDIORAEAEWRIILLRYFNFGAHESQIGEDPRGI.PNNLMPYIQQ
VARAPLEINIVGHDPYDGTGTAIRDYIHVMDLADHIAALKLFTTINIGCTANLGG
TGRGTVLSEWAAFEKASGKPIPKMPPRPDGTAVVASTEKAEKELGWRKAYVEE
MCRDQKWASNNPWGQGRH"

ORIGIN
Query Match 51.3%; Score 673.2; DB 8; Length 1474;
Best Local Similarity 76.8%; Pred. No. 1.3e-152;
Matches 822; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 34 ATGGTGCTCTCCCAACACATCTCGTCCACGGTGGTCCGGTTTCATTGGCACCCAC 93
DB 91 AGGATGGCGTCAGGGGAAACAATCTGGTAATCTGGAGGAGCTGGATTCATCGGATCTAC 150
QY 94 ACCGTGCTTCAGCTTCTCAAGCTGGCTTCAGCGTTTCAATATCGACAATTCGATAC 153
DB 151 ACGGTGCTTCAGCTTCTGAAGCAAGGTTTCAGGTATCCATCATCGACAATCTCTACAC 210
QY 154 TCGGTATGAAGACAGTGGACCGCTCCGCCAAGTGGTGGCCCTGCTTCTCAGAC 213
DB 211 TCGGTATGACGACCGCTCCATAGGGTTCGGCTTTTGGTGGTCCATCTCTCTCCAGCAC 270
QY 214 CTCGAATTCACCCAGGCGCATCTCCGGAATPAGGATGACTTGGAGAAAATCTCTTCCAAA 273
DB 271 CTCGAATTCACCCAGGCGCATCTCCGGAATPAGGATGACTTGGAGAAAATCTCTTCTCAA 330
QY 274 ACAACATTTGATCGGTGATGCTACTTGTGGTGTGAAGCGGTGCTCTGAAGCGTTGGC 333
DB 331 ACCAAATTTGATCGGTGATGCTACTTGTGGTGTGAAGCGGTGCTCTGAAGCGTTGGC 390
QY 334 AAGCCCGTGCCTATTGTGATTTAAATTTGGTGGCACCATCAACCTCTACGAGTTTATG 393
DB 391 AATCCAGCAACTATTATGACACAAATTTGGTGGCACCATCAACCTCTTCCAGTTATG 450
QY 394 GCAAGATTAATTCGAAAAAGATGGTTTCTCATCATCTGCAACCGTTATGGCCAACT 453
DB 451 TCCAAATTCGAAATGAAAAAGCTCGTCATCTCTTCATCTCGGACTGTTTATGGCCAACT 510
QY 454 GAAAGATACCGTGTGAGGAGGATTTCAAGTTCAAGCTATGAATCCCTATGACGAC 513
DB 511 GATCAAAATTCATCGCTTGGAGACTCTAATCTGACGCTATGAATCCCTATGACGAC 570
QY 514 AAGCTTTTCTCGAAGAAATTTGCCGAGATTAATCAGAAAGCTGAACAGAGTGAAGATC 573
DB 571 AAGCTGTTTCTGAGGAAGTAGCACGGGACATTCAGAGGCGCAAGCAGAGTGAAGATC 630
QY 574 ATATTCTGAGATCTTCAATCAGTTGGGCTCATGAAGTGGCAAACTCGTGAAGAT 633
DB 631 ATATTCTGAGATCTTCAATCAGTTGGGCTCATGAAGTGGCAAACTCGTGAAGAT 690
QY 634 CCCAAGGGCATCCCAAAATACCTCATGCTTACATTCAGCAAGTAGCTGTGGAAGATTG 693
DB 691 CCAAGGGCTCCCAAAACACCTCATGCTTACATCAACAAAGTGTGCTGTGCGAGATTG 750
QY 694 ACTGAATCAATGATACGGTCAATGATATTCAGCGGAGTGGCTCTGGATCCGGAC 753
DB 751 CCTGAATCAACATTTATGCTGATGATATCCCACTAAAGATGGCACTGCGATTCCGGAT 810
QY 754 TATATCATGATGAGTGGACTTGGCAGATGGCCATATTGCTGCCCTGCGAAGCTCTTCA 813
DB 811 TATATCATGATGAGTGGACTTGGCAGATGGCTGATGCTGCTGCCCTAGCGAGCTTTTCA 870

```

```

QY 814 ACGGAGAACATAGTGTCTACTGCTTACAACTGGGAACCTGGTCTGGAAACATCTGTCCT 873
DB 871 ACAGACAACATAGTGTCTACTGCTTACAACTGGGAACCTGGTCTGGTACGTCAGTCTT 930
QY 874 GAAATGTTACAGCATTTGAAAAGGCTTCTGGCAAGAAAATTCAGTAAATATATGTCCA 933
DB 931 GAAATGTTGTCAGCATTTGAAAAGGCTTCTGGCAAGAAAATTCAGTAAATATATGTCCA 990
QY 934 AGAAGACCGGAGATGCGACTGAGGTTTATGCTACTACAGAGAGCTGAGAAAACATTT 993
DB 991 AGGAGGCGCAGGGAATGCTACTGCTGTTTATGCTACTACGAGAGGCTGAGAAAACATTT 1050
QY 994 GCTTGGAAAGCAAACTACTGTTGTCGAGAGATGTCAGGACCAATGGAATTTGGGCAAG 1053
DB 1051 GCTTGGAAAGCAAACTACTGTTGTCGAGAGATGTCAGGACCAATGGAATTTGGGCAAG 1110
QY 1054 AACATCCCTGGGCTTACGCGGGAGACCTTGATTTAGCTTGAATAATAT 1103
DB 1111 AATAATCCATGGGGTATCAAGGGAAGCATTTATTTTCTTTCTTATAT 1160

RESULT 3
BT000032 1170 bp mRNA linear PLN 19-SEP-2002
LOCUS Arabidopsis thaliana uridine diphosphate glucose epimerase
DEFINITION (At1g12780) mRNA, complete cds.
ACCESSION BT000032
VERSION BT000032.1 GI:23197647
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1170)
AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shim,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (19-SEP-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arabesequence.stanford.edu

COMMENT
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and
sequenced the pENTR (ORF) clones using the RAFL cDNAs: Nguyen,M.,
Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H.,
Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W.,
Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Toroumi,M.,
Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S.,
Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
Location/Qualifiers
1..1170
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U16095"
/ecotype="Columbia"
source

```



```

/notes="This clone is in pENTR/SD-dTropo. This is a cloned
PCR product using RIKEN clone RAFL05-11-004 (AY120709) as
a template"
1. .1170
/genes="Atlg12780"
/notes="synonym: F13K23.3"
1. .1056
/genes="Atlg12780"
/codon_start=1
/product="uridine diphosphate glucose epimerase"
/protein_id="AAM1531.1"
/db_xref="GI:23197648"
/tranlation="MGSSVEQNILVTGGAGFTGTHVQLLKDGDFKYSIIDNFNSVVI
EAYDRVNIVLGPDLKSLFDNLGDKIDIKLFSKQRFPAVTHFAGLKAVGESVCR
PRAYFNNILVGLDILNLTMAKYNKCMYFSSATVGGPEKIPCEWDFELKAMPYVGR
KLPFLBELARDIKAPSEPNWKRIILLRFYFNPGAHESGISGEDPKGI GNTPMLNLYIQQVAV
GRUPELVNIVGHDIYPTEDGSAVARDYTHVMDLADGHLAALRLFADPKI GCTAYNLGTGQ
TGLPVEWFAAFKAGSKKIPKLCPPRSSDATAVVASTAEKAEKGLMKYGVDEMCEVR
DQKSWANNPFGKQVSKIP"

```

gene

/gene="At1g12780"
/note="synonym: F13K23 3"

CDS
1. 1056
/HOLE="BYHONYM": F13KZ3.3

```

cds
1. .1038
/gene="A"

```

```
/gene=accgiz/oo
/codon start=1
```

```
product="input/
```

```
/product=uridine triphosphate g-proteins
/protein_id="AAN15351.1"
```

```
/protein_id=AA13551.1
/db_xref="GI:23197648"
```

```

/w_xlet=GI:2313/848
/translation="MGSSVEQ

```

71a1ub1a1c1d1e1f1g1h1i1j1k1l1m1n1o1p1q1r1s1t1u1v1w1x1y1z1aa1ab1ac1ad1ae1af1ag1ah1ai1aj1ak1al1am1an1ao1ap1aq1ar1as1at1au1av1aw1ax1ay1az1ba1bb1bc1bd1be1bf1bg1bh1bi1bj1bk1bl1bm1bn1bo1bp1bq1br1bs1bt1bu1bv1bw1bx1by1bz1ca1cb1cc1cd1ce1cf1cg1ch1ci1cj1ck1cl1cm1cn1co1cp1cq1cr1cs1ct1cu1cv1cw1cx1cy1cz1da1db1dc1dd1de1df1dg1dh1di1dj1dk1dl1dm1dn1do1dp1dq1dr1ds1dt1du1dv1dw1dx1dy1dz1ea1eb1ec1ed1ee1ef1eg1eh1ei1ej1ek1el1em1en1eo1ep1eq1er1es1et1eu1ev1ew1ex1ey1ez1fa1fb1fc1fd1fe1ff1fg1fh1fi1fj1fk1fl1fm1fn1fo1fp1fq1fr1fs1ft1fu1fv1fw1fx1fy1fz1ga1gb1gc1gd1ge1gf1gg1gh1gi1gj1gk1gl1gm1gn1go1gp1gq1gr1gs1gt1gu1gv1gw1gx1gy1gz1ha1hb1hc1hd1he1hf1hg1hh1hi1hj1hk1hl1hm1hn1ho1hp1hq1hr1hs1ht1hu1hv1hw1hx1hy1hz1ia1ib1ic1id1ie1if1ig1ih1ii1ij1ik1il1im1in1io1ip1iq1ir1is1it1iu1iv1iw1ix1iy1iz1ja1jb1jc1jd1je1jf1jg1jh1ji1jj1jk1jl1jm1jn1jo1jp1jq1jr1js1jt1ju1jv1jw1jx1jy1jz1ka1kb1kc1kd1ke1kf1kg1kh1ki1kj1kk1kl1km1kn1ko1kp1kq1kr1ks1kt1ku1kv1kw1kx1ky1kz1la1lb1lc1ld1le1lf1lg1lh1li1lj1lk1ll1lm1ln1lo1lp1lq1lr1ls1lt1lu1lv1lw1lx1ly1lz1ma1mb1mc1md1me1mf1mg1mh1mi1mj1mk1ml1mm1mn1mo1mp1mq1mr1ms1mt1mu1mv1mw1mx1my1mz1na1nb1nc1nd1ne1nf1ng1nh1ni1nj1nk1nl1nm1nn1no1np1nq1nr1ns1nt1nu1nv1nw1nx1ny1nz1oa1ob1oc1od1oe1of1og1oh1oi1oj1ok1ol1om1on1oo1op1oq1or1os1ot1ou1ov1ow1ox1oy1oz1pa1pb1pc1pd1pe1pf1pg1ph1pi1pj1pk1pl1pm1pn1po1pp1pq1pr1ps1pt1pu1pv1pw1px1py1pz1qa1qb1qc1qd1qe1qf1qg1qh1qi1qj1qk1ql1qm1qn1qo1qp1qq1qr1qs1qt1qu1qv1qw1qx1qy1qz1ra1rb1rc1rd1re1rf1rg1rh1ri1rj1rk1rl1rm1rn1ro1rp1rq1rr1rs1rt1ru1rv1rw1rx1ry1rz1sa1sb1sc1sd1se1sf1sg1sh1si1sj1sk1sl1sm1sn1so1sp1sq1sr1ss1st1su1sv1sw1sx1sy1sz1ta1tb1tc1td1te1tf1tg1th1ti1tj1tk1tl1tm1tn1to1tp1tq1tr1ts1tt1tu1tv1tw1tx1ty1tz1ua1ub1uc1ud1ue1uf1ug1uh1ui1uj1uk1ul1um1un1uo1up1uq1ur1us1ut1uu1uv1uw1ux1uy1uz1va1vb1vc1vd1ve1vf1vg1vh1vi1vj1vk1vl1vm1vn1vo1vp1vq1vr1vs1vt1vu1vv1vw1vx1vy1vz1wa1wb1wc1wd1we1wf1wg1wh1wi1wj1wk1wl1wm1wn1wo1wp1wq1wr1ws1wt1wu1wv1ww1wx1wy1wz1xa1xb1xc1xd1xe1xf1xg1xh1xi1xj1xk1xl1xm1xn1xo1xp1xq1xr1xs1xt1xu1xv1xw1xx1xy1xz1ya1yb1yc1yd1ye1yf1yg1yh1yi1yj1yk1yl1ym1yn1yo1yp1yq1yr1ys1yt1yu1yv1yw1yx1yy1yz1za1zb1zc1zd1ze1zf1zg1zh1zi1zj1zk1zl1zm1zn1zo1zp1zq1zr1zs1zt1zu1zv1zw1zx1zy1zz

ORIGIN

Query Match	49.5%;	Score 649.2;	DB 8;	Length 1170;
Best Local Similarity	76.3%;	Pred. No. 8.7e-147;		
Matches 798; Conservative	0;	Mismatches 248;	Indels	0;
Gaps	0;			

49	CAA	CAC	AT	CT	TGG	T	CAC	CGG	TGG	TGG	CGG	T	T	CAT	TGG	CA	CC	CA	CG	T	G	T	T	CAG	T	108
19	CAG	A	C	A	T	T	C	T	T	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	78
109	CTC	AA	G	C	T	T	C	A	G	T	T	C	A	A	T	T	C	G	A	T	T	T	C	G	A	168
79	CTC	AA	G	A	G	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	138	
169	GTG	A	C	C	G	G	T	C	C	A	A	G	T	T	T	T	T	T	T	T	T	T	T	T	228	
139	GTT	G	A	T	A	G	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	198	
229	GCG	A	T	C	T	C	C	G	A	T	A	G	A	T	T	T	T	T	T	T	T	T	T	T	288	
199	G	T	G	A	T	C	T	A	A	G	A	A	C	A	T	T	T	T	T	T	T	T	T	T	258	
289	GTG	A	T	C	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	348	
259	GTG	A	T	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	318	
349	T	T	T	G	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	408	
319	T	T	T	G	A	T	A	A	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	378	
409	AAA	A	A	G	A	G	T	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	468	
379	AAA	A	T	G	A	T	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	438	
469	G	A	G	A	G	A	T	T	C	A	G	T	A	T	A	G	A	T	A	G	A	T	T	T	528	
439	A	T	G	A	G	A	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	498	
529	G	A	A	T	T	G	C	C	G	A	T	A	T	T	C	A	G	A	A	T	T	T	T	T	588	
499	G	A	A	T	A	G	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	558		
589	T	T	C	A	T	C	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	648		
559	T	T	C	A	T	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	618		
649	A	A	T	A	C	C	T	C	A	T	T	C	A	G	A	T	T	T	T	T	T	T	T	T	708	
619	A	A	T	A	C	C	T	C	A	T	T	C	A	A	G	T	T	T	T	T	T	T	T	T	678	
709	T	A	C	G	T	C	A	T	A	T	T	C	C	A	G	G	A	T	G	C	T	T	T	T	768	
679	T	A	T	G	A	C	A	T	A	T	T	C	C	A	G	A	T	G	T	A	G	A	T	T	738	

Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES
source
Location/Qualifiers
1..1313
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAPL09-23-D03 (R13909)"
/ecotype="Columbia"
1..1104
105..1160
/note="uridine diphosphate glucose epimerase"
/codon_start=1
/product="At1g12780/F13K23_21"
/protein_id="AA106868.1"
/db_xref="GI:15809880"
/translation="MGSSVSEQNLLVTGGAGFTGTHVVQLLKDGPKFVSIIDNFNSVNI
EAVRVELVGPDLSKLDNLGDLRKGDIKLKSKRFDVAHFAGLKAAGSVEN
PRYFDNNLVLTINLEYWAKYKCKMVFSSATVYGQPEKIPCMEDFELKAMNYPGR
TKLFEIATRDIOKAEPEKPIILLRYFNPVGAHESGICEDEPKGIPNNLMPIQOVAV
GRLEPNAVGHDPTEGDSAVRDYIHVMDLADGHIAALRKLPAQPKIGCTAYNLGTQ
GSVLEWVAPEKASGKKIPIKLPRSRGDATAVYASTEKAKEGLWKAKYGVDEMCR
DQWKANNPNWGYNKL"
1161..1313

3' UTR
ORIGIN
Query Match 49.5%; Score 649.2; DB 8; Length 1313;
Best Local Similarity 76.3%; Pred. No. 8.8e-147;
Matches 798; Conservative 0; Mismatches 248; Indels 0; Gaps 0

QY 49 CAACACATTCCTGGTCA CGGFGTGGCGGTTTCATTGGCACCCACACCGTGGTCAGGTT 108
Db 123 CAGAACATTCCTGGTCTACTGGTGGTGGCTTTATCGGACGCATACACTGTTGTTCAACT 182
QY 109 CTCAAAGCTGGCTCAGCGTTTCAATAATCGACAATTCGATAACTCGGTCAATGAAGCA 168
Db 183 CTCAAAGATGGTTTAAAGTTTCGATCATCGATTAATTTTGATTAATCTGTGTTATCGAGCT 242
QY 169 GTGGACCGCGTCCGCCAAGTGGTTGGCCCTCTCTCTTCTCAGAACCTCAATTCACCCAG 228
Db 243 GTTGATAGATTAGGAGCTTGTGGTCTGCTGATCTCTCCAAAGAGCTGACATCAATCTG 302
QY 229 GCGGATCTCCGGATAGGATGACTTGGAGAACTCTTCTCCAAACACACATTTGATGCC 288
Db 303 GTGTATCTAAGAAACAAGGGAATTGAGAAATCTATTCTCAAGCAGAGATTTGATGCT 362
QY 289 GTGATCCACTTTGTCTGGCTTGAAGCCGGTTGCTGAAAGCGTTGCGAAGCCCGCTCGCTAT 348
Db 363 GTGATTCATTTTGGGGTCTTAAAGCTGTGGGTGAGAGTGTGAAACCCCTCGCGCTAC 422
QY 349 TTGTGATTTAATTTGGTTGGCACCATCAACCTCTAGAGTTTATGCGAAAGTATTAATGTC 408
Db 423 TTTTGAACAATACTGGTTTGAACAATCAATCTATATGAGACCATGGCAAGTACAATGTC 482
QY 409 AAAAAGATGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGGTG 468
Db 483 AAAATGATGGTGTTCATCTTCTGCCACTGTGTTATGGACAACCTGAAAGATTCATGTC 542
QY 469 GAGGAGGATTTCAAGTTTCAAGCTATGAATCCCTATGGAACCGACCAAGCTTTCTGGA 528
Db 543 ATGGAAGACTTTGAATTAAGGCTATGAATCCCTATGGTCGTACTAAGCTCTTTCTTGA 602
QY 529 GAAATGCGCCAGATATTTAGAAAGCTGGAACCAAGATGGAAGATCATATTACTGAGATAC 588
Db 603 GAAATAGCTAGAGATATTTCAAAGCGAAGACCGGAATGGAGAATTAATCTGCTGAGGTAC 662
QY 589 TTCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCA 648

Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Southwick, A. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as FIs.

FEATURES

source

1. .1420
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL05-11-C04 (R16095)"
/ecotype="Columbia"
/note="This clone is in pBluescript"

gene

1. .1420
/gene="At1g12780"

CDS

211. 1266
/note="synonym: F13K23.3"
/gene="At1g12780"

/codon_start=1
/product="uridine diphosphate glucose epimerase"

/protein_id="AA053267.1"

/db_xref="GI:21539429"

/translation="MGSSVEQNILVTGGAGFTGTHVQVLLKDGPKVSIIDNFQNSVI
EAVDRVELVGLDLSKLDLNLGDLRNKDIKLFKORFADVIHFAGLKAVGSVEN
PRYFNNLVGINLYETWAKYKRMFSSSATYVGGPEKIPCKMDEFELKAMPEYGR
TKLLEIARDIOKAEPEWRIILIRYFNVPVGAHESGSGEDPKGIPNNLMPIQQYAV
GRLEPNLVGHDPYEDGSAVDYIHVMDLADGHIALLKLPADPKIGCTAYNLGTGG
GTSLEPMVAAFKASGKIPILKLPFRSGDATAVYASTKAEKELGWKAKYGVDMCR
DQKWANNPNWPGYQNKL"

ORIGIN

49.5%; Score 649.2; DB 8; Length 1420;

Query Match 76.3%; Pred. No. 9e-147; Indels 0; Gaps 0;

Matches 799; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 49 CAACACATTCGTGTCACCGTGTGCGCGTTTCATTGGCACCCACACCGTCTGTCAGCTT 108
DB 229 CAGAACATTCCTTACTGTTGTTGCTGCTGCTTATCGGACGCATCTGTTTCAACTT 288
QY 109 CTCAAAGCTGGTTCAGCTTCAATTAATGCAATTTTCGTAATCTCCGTCTATGGAACA 168
DB 289 CTCAAAGATGTTTAAAGTTTCGATCATGATAATTTGATAACCTCTGTTATCGAAGCT 348
QY 169 GTGACCGCTCGCGCAAGTGTGTCCTCTGCTTCTCAGAACCTCCAAATTCACCCAG 228
DB 349 GTTGATAGATTAGGAGCTTGTGTCCTGATCTCTCCAGAGCTCGACTTCAATCTG 408
QY 229 GCGATCTCCGGAATAGGATGACCTTGGAGAACTCTTCTCCAAACAACTTTGATGCC 288
DB 409 GGTGATCTAAGAAACAAAGGGGACATTTGAGAACTATCTCCAGCAGAGATTGATGCT 468
QY 289 GTGATCCACTTGTGCTTGAAGCGTTGTGCTGAAGCGTTGCGAAGCCCGTCTGCTAT 348
DB 469 GTGATTCATTTGGCGGTCTTAAAGCTGTGGGTGAGAGTTTGAACCCCTCGCGCTAC 528
QY 349 TTTGATTTTAAATTTGGTGGCAATCAACCTCTACAGTTTATGCGAAAGTATATTTGC 408
DB 529 TTTGACAAATTAATTTGGTGGCAATCAACCTCTACAGTTTATGCGAAAGTATATTTGC 588
QY 409 AAAAGATGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGCAAAAGTATACCGTCT 468
DB 589 AAAATGATGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGCAAAAGTATACCGTCT 548
QY 469 GAGGAGGATTTCAAGTTTCAAGTATGAATCCCTATGGAAGCGGACCAAGCTTTTCGGAA 528
DB 649 ATGGAAGACTTTGAATTAAGGCTATGAATCCCTATGAGTCTTAAGCTCTTTCTTGA 708
QY 529 GAATTTCCCGAGTATTCAGAGCTGACACCAAGTATGAGATCATATTTACTGAGATAC 588
DB 709 GAATGCTAGAGATATTCAAAAGGCGAACCAGGAGGAGATTTTCTGCTGAGGTAC 768
QY 589 TTCAATCCAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGCGATCCCA 648

DB 769 TTCAATCTCTAGGACACATGAGCTGGCAGTATTGGTGGAGATCCAAAAGGCATCCCC 828
QY 649 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAGAGTACTGAACCTCAATGTA 708
DB 829 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAGAGTACTGAACCTCAATGTA 888
QY 709 TACGGTCATGATTATCAACGAGGAGTGGCTCTGCGATCGGAGCTATATCCATGTTGATG 768
DB 889 TATGGACATGACTATCCACCGGAGTGGTAGTGGCTTAAGAGCTACATCCATGTTGATG 948
QY 769 GACTTGGCAGATGCGCATATTGCTCGCCTCGCAAGCTTTCACAAAGGAGACATAGGT 828
DB 949 GATTAGCAGATGCGCATATTGCTCGCCTCGCAAGCTTTCACAAAGGAGACATAGGT 1008
QY 829 TGTACTGCTTACAACTGGGAACTGGTCTGTTGGAACATCTGTTGTTGAAATGGTTACAGCA 888
DB 1009 TGTACTGCTTACAACTTAGGGAAGTGGTCAAGGAACGCTCTGTTGTAGAAATGGTTGAGCT 1068
QY 889 TTTGMAAGGCTTCTGCAAGAAATTTCCAGTAAATTTATGTCACAAAGACCGGGAGAT 948
DB 1069 TTTGMAAGGCTTCTGCAAGAAATTTCCAGTAAATTTATGTCACAAAGACCGGGAGAT 1128
QY 949 GCGACTGAGTTTATGATCTCAGAGAGAGCTGAGAGAACTTGGTTGGAAGGCAAC 1008
DB 1129 GCAACAGCAGTTTATGCTTCAACAGAGAGAGCTGAGAGAACTTGGCTGGAAAGGCAAAA 1188
QY 1009 TATGTTGAGGAGATGTCAGGAGCAATGGAATTTGGCAAGAACTTCCCTGGGGT 1068
DB 1189 TATGAGTGGATGAGATGTCAGAGATCAGTGAATTTGGCAAACTATATCCATGAGGT 1248
QY 1069 TACCGGGGAGAGCTTGAATAGCTT 1094
DB 1249 TACCAGATAAGCTTTCGATTTACTT 1274

AX507662 1056 bp DNA linear PAT 27-SEP-2002
Sequence 2357 from Patent WO0216655.
AX507662
AX507662.1 GI:23388899
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsi.

1
Harper, J.P., Kreps, J., Wang, X. and Zhu, T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 2357 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)

Location/Qualifiers
1..1056
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

Query Match 49.1%; Score 644.4; DB 6; Length 1056;
Best Local Similarity 76.3%; Pred. No. 1.3e-145;
Matches 792; Conservative 0; Mismatches 246; Indels 0; Gaps 0;

QY 49 CAACACATTCGTGTCACCGTGTGCGCGTTTCATTGGCACCCACACCGTCTGTCAGCTT 108
DB 19 CAGAACATTCCTTACTGTTGTTGCTGCTTATCGGAGCAGATCTGTTGTTCAACTT 78
QY 109 CTCAAAGCTGGCTTCAAGTTCATTAATCGACAAATTCGATAACTCCGTCATCGGAAGCA 168
DB 79 CTCAAAGATGTTTAAAGTTTCGATCATCGATAATTTGATTAATCTGTTATCGAAGCT 138

169 GTGACCGCGTCCGCAAGTGGTTGGCCCTCTCTCTTCTCAGAACCTCAATTCACCCAG 228
139 GTTGATAGAGTTAGGAGCTTGTGTGCTCTGATCTCTCCAGAAGCTGACATCAATCTG 198
229 GCGGATCTCCGGAATAGGATGACTTGGAGAACCTCTCTCCAAAACACATTTTCATGCC 288
199 GGTGATCTAAGAAACAAAGGAGACATTTGAGAATCTATCTCCAGCAGAGATTTGATGCT 258
289 GTGATCCACTTTGCTGGCTTGAAGCGGTTGTCTGAAAGCGTTTGCAGAGCCCGTCGCTAT 348
259 GTGATTCATTTTGGCGGTCTTAAAGCTGTGGGTGAGAGTGTGAAACCTCGCGCTAC 318
349 TTTGATTTTAAATTTGGTGGCAACCTCAACCTCTAGAGTTTATGCGAAGTATTAATTCG 408
319 TTTGACAAATACCTTGTGGTGGAAACATCAATCTATATGAGACCATGCGAAGTACAACTGC 378
409 AAAAAGATGGTTTCTCATCATCTGCAACCGTTTATGGCCCAACCTGAAAAGATACCGTGT 468
379 AAAATGATGGTGTTCATCTCTGCACTGTTTATGGACAACCTGAAAAGATTCATGTC 438
469 GAGGAGATTTCAAGTTACAGTATGAATCCCTATGAGCGGACCAAGCTTTTCCTGAA 528
439 ATGGAAGACTTTGAATTAAGGCTATGAATCTCTTATGGTCGTAAGCTCTTTCTTGA 498
529 GAAATTTGCCGAGATATTCAGAAAGCTGAACAGCAATGGAATCATATTAATCTAGATAC 588
499 GAAATAGCTAGATATTCRAAGGCGAGAACCGGATGAGATTAATCTGCTGAGGTAC 558
589 TTCAATCCAGTTGGGCTCATGAATGGCAACTCGGTGAAGATCCCAAGGGCATCCCA 648
559 TTCAATCCTGTAGGACACATGAGTGGCGATGATGTTGGTGAAGATCCCAAGGATCCCC 618
649 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAAGATGACTCAACTCAATGTA 708
619 AATACTCATGCTTACATTCACAGTGGCGGTGGACGTTTACCGGAATCAATGTC 678
709 TACGGTCAATGATPATCAACGAGGATGCTCTGCGATCCGGGACTATATCCATGTGATG 768
679 TATGGACATGACTATCCACCGAGGATGCTAGTGGCGGTGAAGACTTACATCCATGTGATG 738
769 GACTTGGCAGATGCGCATATGTCGCCCTCGCAAGCTCTTCCACACGAGAACATAGT 828
739 GATTTAGCAGATGCGCATATGCTGCGCTCAGGAAGCTATTTGCTGATCCAAAGATTTGT 798
829 TGTACTGCTTACAACTGGGAACTGCTCGTGGAAACATCTGTGCTTGAATGTTTACAGCA 888
799 TGTACTGCTTACAACTAGGACTGCTCAAGGAACGCTGTGTTAGAAATGTTGCAGCT 858
889 TTTGAAAAGCTTCTCGCAAGAAAATTCAGTAAATTAATGTTCCAGAGACCGGAGAT 948
859 TTTGAAAAGCTTCTCGCAAGAAAATTCCTCGATTAAGCTCTGTCGCAAGAGGTCAGGAGAT 918
949 GCGACTGAGGTTTATGATCTCAGAGAGAGCTGAGAAAAGACTGTTGGGAAGGCAAC 1008
919 GCAACGACGATTTATGCTTCAACAGAGAGGCTGAGAAAGACTTGGCTGGGAAGCAAA 978
1009 TATGTTGTGAGAGATGTCAGGGAACCAATGGAATTTGGGCAAGAACCAATCCCTGGGCT 1068
979 TATGAGTGGATGAGATGTCAGAGATCAGTGAATGGAATGGGCAACCAATTAATCCATGGGCT 1038
1069 TACCGGGGAAGCTTGA 1086
1039 TACCAGAAATAGCTTGA 1056

RESULT 7
AF334724
LOCUS AF334724 1056 bp mRNA linear PLN 07-MAY-2003
DEFINITION Arabidopsis thaliana clone C00128 (e) putative uridine diphosphate
glucose epimerase (Atg12780) mRNA, complete cds.
ACCESSION AF334724
VERSION AF334724.1 GI:12248020

KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L.,
Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R.,
Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A.,
Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 1056)
AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
Goldsmith,A.D., Lee,J.M., Toriumi,M., Yu,G., Brooks,S., Chao,Q.,
Chen,H., Karlin-Neumann,G., Kim,C., Lam,B., Miranda,M., Nguyen,M.,
Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and
Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
REFERENCE 3 (bases 1 to 1056)
AUTHORS Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L.,
Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R.,
Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A.,
Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2003) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
COMMENT Annotation based on July 2002 version of the Arabidopsis genome
submitted to Genbank.
FEATURES
source
1..1056
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="C00128 (e)"
/ecotype="Columbia"
/note="This clone is now in pUNI 51 vector. Previously,
it was in pUNI-T3-D/V5-His-TOPO under the clone number
C00128."
gene
1..1056
/gene="at-g12780"
CDS
1..1056
/gene="atg12780"
/codon_start=1
/evidence=experimental
/product="putative uridine diphosphate glucose epimerase"
/protein_id="AA050102.1"
/db_xref="GI:12248021"
/translation="MGSSVEQNILVTGGAGFIGTHVVQLLKDFKVSIIIDNPDNSVI
EAYDRVELVGLDLSKLDLNLKNGDIEKLFQKQRFDAVIHFAGLKAQVESVEN
PRYFDNLVGTINLVETAKYKRWVFSSATVYQPEKIFCMEDFELKAMNYPGR
TKLFLEIARDIOKARPEWRIILIRYFNPVAGSGSIGDEPKGIPNNLMPYIQQAV
GRPELVNYGHDYPTDGSVARDYIFVMDLADGHIALLKRLFADPKIGCTAYNLGTQ
GTSVLENVAAFEKASCKKIPKLPKRRSGDATAVASTEKAKELGWKAKYGVDMCR
DQWKAANNPNWGTQNKL"
ORIGIN
Query Match 49.1%; Score 644.4; DB 8; Length 1056;
Best Local Similarity 76.3%; Pred. No. 1.3e-145;
Matches 792; Conservative 0; Mismatches 246; Indels 0; Gaps 0;
QY 49 CAACACATTTGTCACCGTGGTGGCGGTTTCAATGGCACCCACCGTCTGACGTT 108
Db 19 CAGAACATTTCTGTACTGTGGTGGCTTTATCGGGACGCATCTGTTTCAACTT 78
QY 109 CTCAAAGCTGCTTACCGTTTCAATATCGACAATTTTCGATAACTCCGTCATCGGAAGCA 168
Db 79 CTCAAAGATGTTTAAAGTTTCATCATCGATAATTTTGATAACTCTGTTATCGAGCT 138


```

Db      949  TGTACTGCTTCAATCTAGGACGTGTCAAGGAACGCTGTGTAGAAATGGTTGCAGCT 1008
Qy      889  TTTGAAAGGCTTCTGGCAAGAAATTCAGTAAATATATCTCCAAAGAACCGGGAGAT 948
Db      1009  TTTGAAAGGCTTCCGGCAAGAAATCCGATTAAGCTCTCTCCGAGAGTCAAGGAGAT 1068
Qy      949  GCGACTGAGGTTATGCACTCTACAGAGAGACTGAGAAAGAACTTGGTTGGAAGGCAAC 1008
Db      1069  GCAACAGCAGTTTATGCTTCAACAGAGAGGCTGAGAAAGAACTTGGCTGGAAGGCAAA 1128
Qy      1009  TATGCTGTGGAGGAGATGTGACGAGGACCAATGGAAATGGGCAAGAAACAATCCCTGGGGT 1068
Db      1129  TATGAGTGGATGAGATGTGAGAGATCAGTGGAAATGGGCTTTCAATATCCATGGGT 1188
Qy      1059  TACGCGGGAGCCTTGAATTAGCTT 1094
Db      1189  TACCAGAATAAGCTTGAATTACTT 1214

RESULT 9
AK117913
LOCUS      1263 bp mRNA linear PLN 06-DEC-2002
DEFINITION Arabidopsis.thaliana At1g63180 mRNA for putative uridine
            diphosphate glucose epimerase, complete cds, clone: RAF119-09-L10.
ACCESSION AK117913
VERSION    AK117913.1 GI:26450887
KEYWORDS   FLI CDNA; CAP trapper.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1
AUTHORS    Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
            Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
            Kawai, J., Hayashizaki, Y., and Shinozaki, K.
TITLE      Arabidopsis thaliana full-length cDNA
JOURNAL    Published Only in Database (2002)
REFERENCE 2 (bases 1 to 1263)
AUTHORS    Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
            Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
            Kawai, J., Hayashizaki, Y., and Shinozaki, K.
TITLE      Direct Submission
JOURNAL    Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
            Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
            230-0045, Japan (E-mail: mseki@gsc.riken.go.jp,
            URL: http://pfweb.gsc.riken.go.jp. Tel: 81-45-503-9625,
            Fax: 81-45-503-9586)
COMMENT    An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al. (1998) plant J. 15:707-720;
            Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et
            al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
            This clone is in a modified pluescript vector.
            Please visit our web site (http://pfweb.gsc.riken.go.jp/) for
            further details.
FEATURES   Location/Qualifiers
            1..1263
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /cultivar="Columbia"
                /db_xref="taxon:3702"
                /chromosome="1"
                /clone="RAF119-09-L10"
                /note="common name: thale cress"
            1..1263
                /gene="At1g63180/F16M19_8"
                /codon_start=1
                /product="putative uridine diphosphate glucose epimerase"
                /protein_id="BAC42551.1"

```

```

/db_xref="GI:26450888"
/translation="MGSSVEQNILVTGGAGFIGTHTVVLNQLNQKFKVLIIDLNSV
EAVHYRELVDGLSTKLEMAKNGDLEKFSNQRFDVIFHFAKKAQVESVGN
PEPRYDNLVGTINLYETMAKNCOMYFSSATVYQPEIVPCVDFELQAMNBYGR
TKLFLERARDTHAAEPKWKIILLFYFVGAHESGRIGEDPKGI:PNLMPYIQVAV
GRLPLNVGHDYPTMDGSVRDYIHWMDLADGHVAALNKLKIFSDSKIGTANLGTQ
GTSVLEWYSSPKASGKKIPIKLPFRAGDATAVAVASTQKAEKELGKAKYGVDMCR
DOWNWANKNPWFQKPK"

ORIGIN
Query Match      48.1%; Score 631; DB 8; Length 1263;
Best Local Similarity 75.2%; Pred. No. 2.3e-142;
Matches 787; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

Qy      49  CAACACATCTCGTCACCGGTGTCGCGTTCATTGGCACCCACACCGTCGTTCCAGCTT 108
Db      70  CAGAACATTTTGGTGACTGGTGTGGATTCATTGGNACACATCTGTTGTTGAGCTT 129
Qy      109  CTCAAAGCTGGCTTCAGCGTTTCAATAATCGACAATTTGATAACTCCGTTGAAGCT 168
Db      130  TTGATACAGGGTTTAAAGGTTACGATCATTTGATAATCTCTGTCGTTGAAGCT 189
Qy      169  GTGGACCGGTGCGCCAGTGGTGGCCTCTGCTTCTCAGAACTCCCAATTCACCCAG 228
Db      190  GTTCATAGGGTAGGGAATCTGTGTGCTCTGATCTCTTACCAGCTTGAATTCATCTG 249
Qy      229  GCGCATCTCCGGAATAGGGATGACTTGGAGAACTCTTCTCCAAAACAACATTTGATGC 288
Db      250  GGTGATCTAAGAAACAAGAGAGATATTGAGAACTCTTTTCCAAATCAGAGATTTGATGT 309
Qy      289  GTGATCAGTTTGTGCTTGAAGCGTGTCTGAAAGGTTGCGAAGCGTTCGCGTAT 348
Db      310  GTGATTCATTTGCTGCTCTTAAAGCTGGGAGAAAGTGTGGAAACCTCTGTCGTTAC 369
Qy      349  TTTGATTTAAATTTGGTGGCCACCTCAACCTCTACGAGTTTATGSCAAAGTATATTGC 408
Db      370  TTTGATAATACTTAGTTGGAATCTATATGAGACCATGSCAAATATCAACTGC 429
Qy      409  AAAAGATGGTTTCTCATCTGCAACCGTTTATGGCAACCTGAAAGATACCGTGT 468
Db      430  AAAATGATGGTGTTCGTCGTCTGCAACAGTTTATGGTCAACCTGAAATAGTCCCATGT 489
Qy      469  GAGGAGATTTCAAGTTTACAAGCTATGAATCCTATGAGCGGACCAAGCTTTTCTGTGAA 528
Db      490  GTGGAAGACTTTGAGTTACAGCTATGAATCTTATGGTCGTACTAAGCTGTTCTTGAA 549
Qy      529  GAAATGCCGAGATATTGAGAAAGCTGAACAGAGATGGAAGATCATATTACTGATAC 588
Db      550  GAGATAGTAGAGACATTCACGCTGCGGAACCAAGATGGAAGATAATTTCTCTGAGGTAC 609
Qy      589  TTCATCCAGTTGGGGCTCATGAAAGTGCACAACTCGGTGAAGATCCCAAGGCGCATCCCA 648
Db      610  TTCATCCGGTTGAGCTCAGAGAGTGAAGATTTGGAAGATCCCAAGGCGCATACCG 669
Qy      649  AATAACCTCATGCTTACATTCAGCAAGTAGTGTGGAAAGATTTGACTGAATCATATGA 708
Db      670  AATAATCTCATGCTTATATATCAACAGTGGCGGTGGAAGATTACTGTAATCATATGA 729
Qy      709  TACGCTCATGATTATCCAGCGGGATGCTCTGGATCCGGGACTATATCCATGTGATG 768
Db      730  TTTGGACATGATTATCTTACCATTGATGTAGCGGGTTAGAGACTATCATTCATGTAATG 789
Qy      769  GACTTGGCAGATGGCCATATTGTCGCCCTGCGAAAGCTCTTCAACCGGAGAACATAGGT 828
Db      790  GATTTAGCAGATGGCCATGTGGCTGGTTAAACAATTTGTTTCGACTCAAAGATTGGC 849
Qy      829  TGTACTGCTTACAACTGGGACTGTCGTGGAACATCTGTGCTTGAATTTGGTTACAGA 888
Db      850  TGTACTGCTTATAATCTTGCACTGTCAGGAAGACTCTGTCTTGAATTTGGTTCTTCT 909
Qy      889  TTTGAAAAGGCTTCTGGCAAGAAAATTCAGTAAAAATTATGTCCAAAGAGACCGGAGAT 948
Db      910  TTTGAAAAGGCTTCTGGCAAGAAAATACCTATCAAGCTATGTCCAAAGAGAGCTGAGAT 969

```



```
QY 949 GCGACTGAGTGTATGCACTACAGAGAGCTGAGAAAGAACTTGGTGGAGGCAAC 1008
Db 970 GCAACAGCTGTTTATGCTTCAACTGAGAAAGCTGAGAGAACTTGGCTGGAGGCAAA 1029
QY 1009 TATGGTGTGAGAGAGATGTGACGGGACCAATGGAATTGGGCAAGAACCAATCCCTGGGT 1068
Db 1030 TATGGAGTTGATGAGATGTGACAGATCAATGGAATGCGCAATTAAGAAATCCATGGGT 1089
QY 1069 TAGCGGGGAGCTTGAATCTG 1095
Db 1090 TTCAGAAGAAGCCTTGATGATTG 1116

RESULT 10
AY085887
LOCUS AY085887 1397 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 1932 mRNA, complete sequence.
ACCESSION AY085887
VERSION AY085887.1 GI:21404597
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1397)
AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)
JOURNAL MEDLINE
PUBMED 22088475
REFERENCE 12093376
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1397)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-WAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
GenSet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
FEATURES
Location/Qualifiers
1..1397
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="1932"
197..1252
/codon_start=1
/product="uridine diphosphate glucose epimerase, putative"
/protein_id="AAM63099.1"
/db_xref="GI:21554018"
/translation="MGSSVEQNILVTGGAGFIGTHTVQLNQGGFKVITIDLNSV"

CDS
```

ORIGIN

Query Match 48.0%; Score 629.4; DB 8; Length 1397;

Best Local Similarity 75.1%; Pred. No. 5.8e-142; Indels 0; Gaps 0;
Matches 786; Conservative 0; Mismatches 261;

```
QY 49 CAACACATTTGTTGTCACCGGTGGTCCGGTTTCATTGGCACCCACACCGTGTGAGCTT 108
Db 215 CAGAATATTTTGGTGACTGGTGGTCTGATTCATTGGAAACACATCTGTTGTCAGCTT 274
QY 109 CTCAAAGCTGCTTCAGCGTTTCATATCGACAAATTCGATACCTCCGTCATCGAGACA 168
Db 275 TTGAATCAGGGTTTTTAAGGTTACGATCATGTATATCTTGTATAACTCTGCTGTTGAAGCT 334
QY 169 GTGACCGCGTCCGCAAGTGGTGGCCCTCTGCTTTCTCAGAACCTCCAAATTCACCCAG 228
Db 335 GTTCTAGGGTTAGGGAACCTTGTGGTCTGATCTCTCTACCAAGCTTGAATTCATCTG 394
QY 229 GCGCATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAAACAACATTTGATGCC 288
Db 395 GGTGATCTTAAGAAAACAAGGAGATATTGAGAACTCTTTTCCAAATCAGAGATTTGATGCT 454
QY 289 GTGATCCACTTTGCTGGCTTGAAGCGTGTCTGAAAGCGTTGCGAAGCGCCCGTCGCTAT 348
Db 455 GTGATTCACTTTGTGCTTAAAGCTGTGGAGAAAGTGTGGAAACCCCTCGTCGTAC 514
QY 349 TTTGATTTAATTTGGTGGCACCATCAACTCTACGAGTTTATGGCAAGATATAATGTC 408
Db 515 TTTGATAATAACTTAGTTGGAACTATCAATCTATATGAGACCATGGCAAAATACAACTGC 574
QY 409 AAAAGAGATGTTTTCTCATCATCTGCAACCGTTTTATGCCAACCTGAAAGATACCGTGT 468
Db 575 AAAATGATGGTGTGTTTGTGCTCTGCAACAGTTTATGTCACCTGAAATAGTCCCATGT 634
QY 469 GAGGAGGATTTCAAGTTTACAGCTATGAATCCCTATGAGCGGACCAAGCTTTCTCTGAA 528
Db 635 GTGGAAGACTTTGAGTTACAGGCTATGAATCCTTTATGTCGTACTAAGCTGTGTTCTTGA 694
QY 529 GAAATGCCAGATATTTCAGAAAGCTGAACCAAGTGGAGATCATATTACTCAGATAC 588
Db 695 GAGATAGTAGAGACATTCACGCTGCGGAACCAAGATGGAGATTAATCTCTGAGGTAC 754
QY 589 TTCAATCCAGTTGGGGTTCATGAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCA 648
Db 755 TTCAATCCGGTTGGAGCTCACGAGATGGAAGAAATTCGAGAAGATCCAAAGGGCATACCG 814
QY 649 AATAACCTCATGCTTACATTCACCAAGTAGCTCTGCGAAGATTGACTGAACCTCAATGTA 708
Db 815 AATAATCTCATGCTTATATCCAAAGTGGCGGTGGAGATTACCTGAACCTCAATGTA 874
QY 709 TACGGTCATGATTATCCAAAGGAGATGGCTCTCGGATCCGGGACTATATCCATGTGATG 768
Db 875 TTTGGACATGATTATCCTACCATGATGGTAGCGGGTTAGAGACTACATCCATGTAATG 934
QY 769 GACTTGGCAGATGGCCATATTGCTGCCCTCGCAAGCTCTTCACACGGAGAACATAGGT 828
Db 935 GATTTAGCAGATGGCCATGTGGCTGCGGTAAACAAATTTGTTTTTTCAGACTCAAGATTTGGC 994
QY 829 TGTAATGCTTACAACTCCGGGAACCTGGTCTGTGGAACATCTGTGCTTGAATGTGTACAGCA 888
Db 995 TGTAATGCTTATATCTTGGCACTGGTTCAGGAACCTCTGTCTTAGAATGTGTTTCTTCT 1054
QY 889 TTTGAAAGAGCTTCTGGCAAGAAAATTCAGTAAATATGTTCCAGAGAGACCGGAGAT 948
Db 1055 TTTGAAAGAGCTTCTGGCAGAAAATATACCTATCAAGCTATGTCACAGAGAGAGCTGGAGAT 1114
QY 949 CGCACTGAGGTTTATGCACTACAGAGAGCTGAGAAAGAACTTGGTTGGAAGGCAAC 1008
```


Db 1115 GCACAGCTGTTATGCTTCACTCAGAAAGCTGAGAGAACTTGGCTGGAGCAAAA 1174
 QY 1009 TATGCTGTGGAGAGATGTCGGGACCAATGAATGGTGGCAAGAAACAATCCCTGGGGT 1068
 Db 1175 TATGAGATTTGATGAGATGTCGAGATCAATGGAATCGGCAATGAAGAAATCATGGGGT 1234
 QY 1069 TAGCGGGGAAGCCTTGAATTAGCTTG 1095
 Db 1235 TTCCAGAAAGCCCTGATGATTG 1261

RESULT 11

AY094481
 LOCUS Arabidopsis thaliana Atlg63180/F16M19_8 mRNA linear PLN 05-MAY-2002

AY094481
 DEFINITION Arabidopsis thaliana

AY094481
 ACCESSION AY094481

AY094481.1 GI:20453217

FLI CDNA.

KEYWORDS Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1240)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,

Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,

Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,

Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,

Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,

Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,

Theologis, A., and Ecker, J.R.

Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 1240)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,

Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,

Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,

Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,

Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,

Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,

Theologis, A., and Ecker, J.R.

Direct Submission

Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory

(SIGNAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

There is no suitable ORF in this cDNA clone.

RIKEN Genomic Sciences Center (GSC) members carried out the

sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,

Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Chang, E.,

Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B.,

Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,

Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,

Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,

Theologis, A., and Ecker, J.R.

The Salk, Stanford, P0EC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,

Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Chang, E.,

Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B.,

Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,

Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,

Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,

Theologis, A., and Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)

contributed equally to this work as PIs.

Location/Qualifiers

1. 1240

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

FEATURES

source

/chromosome="11"
 /clone="RAFLO9-87-P16 (R21984)"
 /ecotype="Columbia"

ORIGIN

Query Match 47.4%; Score 621.6; DB 8; Length 1240;
 Best Local Similarity 75.2%; Pred. No. 4.5e-140;
 Matches 788; Conservative 0; Mismatches 259; Indels 1; Gaps 1;

QY 49 CAACACATCTCTGGTCACCGGTGGTGGCTTTCATTGGGACCCACACACCGTGGTTCAGCTT 108
 Db 38 CAGAACATTTTGGTGAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 97
 QY 109 CTCAAGCTGGCTTCAGCGCTTCAATAATCGACAAATTCGATAACTCCGTCATCGAAGCA 168
 Db 98 TTGAATCAGGGTTTAAAGGTTACGATCAATTAATCTTGATAACTCTGTCGTTGAAGCT 157
 QY 169 GTGACCGCTCGCCCAAGTGGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 228
 Db 158 GTTCATAGGGTTAGGGAACCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 217
 QY 229 GGGATCTCCGGAATAGGGATGACTTGGGAGAACTCTTCTCCAAAACAAATTTGATGCC 288
 Db 218 GTGATCTTAAGAAACAAAGGAGATATGAGAACTCTTTCATCAGAGATTTGATGCT 277
 QY 289 GTGATCCACTTGTGGCTTGAAGCGTGTGTAAGCGTGTGTAAGCGTGTGTAAGCGTGTGTAAG 348
 Db 278 GTGATTCACCTTGTGGCTTAAAGCTGTGGGAGAAAGTGTGGGAAACCCCTCGTGGTAC 337
 QY 349 TTTGATTTAAATTTGGTTGGCCACCATCAACTCTACAGTGTATGGCAAAAGTATAATTCG 408
 Db 338 TTTGATTAATACTTAGTTGGAATCTATCATATGAGACCATGGCAAAATCAACTGC 397
 QY 409 AAAAAGATGGT-TTCTCATCTCTGCAACCGTTTATGGCCAACTTGAAGATATACCGTG 467
 Db 398 AAAATGATGGTGTGTTTTCGTCGTCTGCAACAGTTTATGGTCAACCTTGAATAGTCCATG 457
 QY 468 TGAGGAGGATTTCAAGTTTCAAGCTATGAATCCCTATGGACGGACCAAGCTTTTCTCTGGA 527
 Db 458 TGTGGAAGATTTGAGTTTCAAGCTATGAATCCCTATGGTGGTGGTGGTGGTGGTGGTGGT 517
 QY 528 AGAATTTCCCGAGATATTGAGAAAGTGAACCAAGATGAAGATGAAATATATATGAGATA 587
 Db 518 AGAGTAGCTAGAGACATTCACGCTCGGAAACCAAGATGGAAGATTAATTTCTTGAGGTA 577
 QY 588 CTTCAATCCAGTTGGGCTCATGAAAGTGCAACTCGGTGGAAGATCCCAAGGCGCATCCC 647
 Db 578 CTTCAATCCGTTGGAGCTCACGAGAGTGAAGATTTGGAAGATCCCAAGGCGCATACC 637
 QY 648 AAATAAACCCTCATGCTTACATTCAGCAAGTAGCTGTTGGAAGATTCGACTGAATCAATGT 707
 Db 638 GAATAATCTCATGCTTATATCCAAAGTGGCGGTGGAAGATTACCTGAATCAATGT 697
 QY 708 ATAGGTCATGATTATCCACGAGGAGTGGCTCTGGATCCGGGACTATATCCATGAT 767
 Db 698 ATTTGGAATGATTTATCTTACCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 757
 QY 768 GGACTTGGCAGATGGCCATATTCCTCCCTCGCAAGAGCTCTTCACAAACGAGAACATAGG 827
 Db 758 GGATTAGCAGATGGCCATGCTGGCTGGTTAAACAAATTTGTTTCAGACTCAAGATTGG 817
 QY 828 TTGTACTGCTTACAACTGCGGAACTGGTCTGTGGAACATCTGTGCTGTGAAATGTTACAGC 887
 Db 818 CTGTACTGCTTATAATCTTGGCACTGGTCAAGGAATCTCTGCTTGAATAATGTTTCTTC 877
 QY 888 ATTTGAAAGGCTTCTGGCAAGAAATTTCCAGTAAATATGTTCCAAAGAGACCGGAGA 947
 Db 878 TTTTGAAGAGCTTCTGGCAAGAAATACCTATCAAGCTATGTTCCAAAGAGAGCTGGAGA 937
 QY 948 TGGCACTGAGGTTATGCAATCTACAGAGAGCTGAGAAAGAACTTGGTGGAGGCAAA 1007
 Db 938 TGCACAGCTGTTTATGCTTCACTCAGAAAGCTGAGAGAGAACTTGGTGGAGGCAAA 997

QY 534 TGGCCGAGATATTCAGAAAGCTGAACAGAAATGGAAGATCATATTAATCTGAGATCTTCAA 593
 Db 602 CTTCCGCGAGGTTCAGCGCCGCCACCCGGAGATGAGGGTGATCTCTCAGTACTTCAA 661
 QY 594 TCCAGTGGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCAGGGCATCCCAAATAA 653
 Db 662 CCCATCGGGCTCACCGGAGCGGCACATCGGGGAGACCCAGGGGATCCCCACAA 721
 QY 654 CTTATGCCCTTACATTCAGCAAGTAGCTGTGGAAGATGACTGAATCAATGTTATCGG 713
 Db 722 CTTCTTCCGTATCATCCAGCAGGTCCGCTCGGCCGCCGCCAGCTCAAGCTCTACGG 781
 QY 714 TCATGATTATCCAAACAGGATGGCTCTCGATCCGGGACTATATCCATGTGATGACTT 773
 Db 782 CTTCTGACTACCCAAACAGGAGCGGCACCGGATCATGAGGATTACATACATGTTAGTGC 841
 QY 774 GGCAGATGGCCATTTGCTGCCCTCGCAAGCTCTTCAACAGGAGAACATAGTTGTAC 833
 Db 842 TGGCGATGGCCATTTGCTGCCCACTGGAGAGCTCTTCGTAATCTCCCTGACATTTGTTGT 901
 QY 834 TGCTTCAACCTGGGAACCTGCTGCGAATCTGCTGAAATCTGCTGAAATCTGCTGAAATCTG 893
 Db 902 GCGTTACATCTAGGAACAGGCTGTGGAACAGCGTCTCGAGGTGGTGAAGCGTTGCA 961
 QY 894 AAGGCTTTCGCAAGAAATTCACGTAAATTTATGTCCAAGAACCGGGAGATCGAC 953
 Db 962 GGAGGCGTCCGGAAGAAATTCCTATCAAGATTTGCCCCAGAGACCTGGAGATTGCAC 1021
 QY 954 TGAGCTTATGCTATCAGAGAGAGCTGAGAGAGCTGTTGGAGGCAAACTATGG 1013
 Db 1022 TGAGGTTTACGCTTCACATGACAGGCGCAAGGAGGCTCGGATGAGTCTCGGTTGG 1081
 QY 1014 TGTGGAGGAGATGTGAGGAGCAATGGAATGGGCAAGAACAACTCCCTGGGGTTACGC 1073
 Db 1082 AATAGAGGACATGTGAGGAGCAATGGAATGGGCAAGAACAACTCCCTGGGATACAG 1141
 QY 1074 GCGGAGCCT 1083
 Db 1142 CGCCAATGCT 1151

RESULT 13
 AK073610
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033049118, full insert sequence.
 AK073610
 ACCESSION
 VERSION AK073610.1 GI:32983633
 KEYWORDS FLI CDNA; CAP trapper.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
 1 The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team: Kikuchi S., Sato H., Kikuchi S., Sato H., Negata T., Kawagashira N., Doi K., Kishimoto N., Yazaki J., Kodama T., Kurosaki T., Kusumegi T., Lu M., Masuda H., Miura J., Kodama T., Kurosaki T., Niikura J., Oka M., Ryu R., Sugano S., Sugiyama A., Suzuki Y., Tsunoda Y., Ueda M., Xie Q., Yokomizo S., Yoshimura A., Matsubara K. and Murakami K.
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S., Hanagaki T., Hara A., Hashizume M., Hayaehida K., Hayatsu N., Hiramoto K., Ichi M., Kagawa I., Kanagawa S., Katoh H., Kawai J., Kouda M., Kishikawa-Hirozane T., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numasaka R., Ohno M., Osato N., Oca Y., Saich H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Sato K., Shibata K., Shinagawa A., Shiraki T., Sugabe Y., Tagami M., Tagami-Takeda Y., Tagawa A., Takahashi F., Takaku-Akahira S., Tanaka T., Tomaru A., Toya T., Waki K., Yasunishi A. and Hayashizaki Y.
 Location/Qualifiers
 1. 1425
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J033049118"

FEATURES
 source
 Query Match 36.8%; Score 482.8; DB 8; Length 1425;
 Best Local Similarity 66.8%; Pred. No. 2.4e-106;
 Matches 688; Conservative 0; Mismatches 342; Indels 0; Gaps 0;

ORIGIN

japonica rice
 Science 301 (5631), 376-379 (2003)
 22752273
 12869764
 2 (bases 1 to 1425)
 AUTHORS
 Fujimura T., Fukuda S., Hanagaki T., Hara A., Hashizume W., Hayaehida K., Hayaehida K., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Hori F., Iida J., Iida Y., Ikeda R., Imamura K., Kagawa I., Inotani K., Ishibiki J., Ishii Y., Ishikawa M., Itoh M., Kawai J., Kanagawa S., Katoh H., Kawagashira N., Kawai J., Kawamata M., Kikuchi S., Kishikawa-Hirozane T., Kishimoto N., Kobayashi M., Kodama T., Kojima K., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Kurosaki T., Kusumegi T., Li C., Lu M., Masuda H., Matsubara K., Matsuyama T., Miura J., Miyazaki A., Mizuno K., Murakami K., Murata M., Nagata T., Nakamura M., Nishi K., Niikura J., Nishi K., Nomura K., Numasaka R., Numasaka R., Ohno M., Ohtsuki K., Oka M., Oka H., Osato N., Oca Y., Ohtsuki K., Ryu R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Sato K., Sato K., Sato K., Shibata K., Shinagawa A., Shiraki T., Shishiki T., Shishiki T., Sogabe Y., Sugano S., Sugiyama A., Suzuki K., Suzuki Y., Tagami M., Tagami-Takeda Y., Tagawa A., Takahashi F., Takaku-Akahira S., Tanaka T., Tomaru A., Toya T., Tsunoda Y., Ueda M., Waki K., Xie Q., Yanagi W., Yazaki J., Yoshimura A., Yamamoto M., Yasunishi A., Yazaki J., Yokomizo S. and Yamamoto M.
 Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp)
 Tel:81-29-838-7007, Fax:81-29-838-7007
 This clone is one of the 28K full-length cDNA clones from japonica rice.
 URL : http://cdna01.dna.affrc.go.jp/cdna/
 NIAS Rice Full-length cDNA Project Team: Kikuchi S., Sato H., Negata T., Kawagashira N., Doi K., Kishimoto N., Yazaki J., Ishikawa M., Yamada H., Ooka H., Hotta I., Kojima K., Namiki T., Ohneda E., Yanagi W., Suzuki K., Li C., Ohtsuki K., Shishiki T., and Yamamoto M.
 FAIS Genome Sequencing & Analysis Group: Ohtsuki Y., Iida Y., Fujimura T., Ikehara R., Kawamata M., Kobayashi M., Kodama T., Kurosaki T., Kusumegi T., Lu M., Masuda H., Miura J., Kodama T., Kurosaki T., Niikura J., Oka M., Ryu R., Sugano S., Sugiyama A., Suzuki Y., Tsunoda Y., Ueda M., Xie Q., Yokomizo S., Yoshimura A., Matsubara K. and Murakami K.
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S., Hanagaki T., Hara A., Hashizume M., Hayaehida K., Hayatsu N., Hiramoto K., Ichi M., Kagawa I., Kanagawa S., Katoh H., Kawai J., Kouda M., Kishikawa-Hirozane T., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numasaka R., Ohno M., Osato N., Oca Y., Saich H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Sato K., Shibata K., Shinagawa A., Shiraki T., Sugabe Y., Tagami M., Tagami-Takeda Y., Tagawa A., Takahashi F., Takaku-Akahira S., Tanaka T., Tomaru A., Toya T., Waki K., Yasunishi A. and Hayashizaki Y.
 Location/Qualifiers
 1. 1425
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J033049118"

QY 54 CATTCTGTGTCACCGGTGGTGGCGTTTCATTTCGGACCCACACCGTGGTTCAGTCTCTCAA 113
Db 178 CGTCTGTGTGACGGGGGAGCGGGGTACATCGGACGACACCGTGTCTCGGCTCTCTGGA 237
QY 114 AGCTGGCTTCAGCGTTTCAATATGACAAATTCGATACTCCGTCATCGGACGAGTCTGA 173
Db 238 GAAGGGTTCGGGTACCGTCTGTGACAACTTCACAACTCCGTCGCGGAGGCGCTCGA 297
QY 174 CCGGTCGCCCAAGTGGTGGCGTCTGTCTTCAGAACCTCCAAATTCACCCAGGGCGA 233
Db 298 CCGGTCGCCCTCATCGCGGGCGCGCTCTCCGCGCGCTTCGATCTCATCGCGGGGA 357
QY 234 TCTCCGGAATAGGATGACTTCGAGAACTCTTCCTCAAAACAACTTGTGTCGCTGAT 293
Db 358 TCTCAAGACAGGACGACATGAGAGGTGTCGCGCCAGAGGATGACCGCGTAT 417
QY 294 CCACTTTGCTGGCTTGAAGCGGTGCTGAAAGCGTTGCGAAGCCCGCTCGCTATTTTGA 353
Db 418 CCACTTCGCGGCTGAAGCGGTGGGAGAGCGTTCGCGCACCCCGCAGATGACTACGA 477
QY 354 TTTTAATTTGGTGGACCACTCACTCTACGATTTATGCGCAAGTATATGCGAATAA 413
Db 478 GAACAAGCTCGCGGACCATGAACTCTCTCCGCTATGACCAAGTACGCGTGCAGAA 537
QY 414 GATGTTTTCTCATCTCTGCAACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGGA 473
Db 538 GATAGTGTCTCGTCTCGCGACGCTGTACGCGCAGCGGAGAGACCCCTCGCTCGA 597
QY 474 GATTTCAAGTTACAAGCTATGATCCCTATGACGAGACCAAGCTTTCTCGAAGAAT 533
Db 598 GGATTTCAAGCTGAGCGCTCTCAACCCATACGCGACCAACAGCTCGTCTCGGAACTA 657
QY 534 TGCCCGAGATATTCAGAAAGCTGAACAGATGGAAGATCATATTAATGAGATATCTCAA 593
Db 658 CTTCGCGAGGTGACGCGCGCCGACCGGAGATGAGGTGATCTCTCTCAGTACTTCA 717
QY 594 TCAGTTGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCCAAGGAGATCCCAATAA 653
Db 718 CCCATCGCGCTCACCGGAGCGGACATCGGAGGAGGCCCGGAGGAGATCCCAACAA 777
QY 654 CTTATGCTTACATTCAGCAAGTACGTTCTGGAAGATGACTGAACTCAATGATACGG 713
Db 778 CTTCTTCGTATCATCAGCAGGTGCGCTCGCGCCCGCCCGAGCTCAAGTCTACGG 837
QY 714 TCATGATATTCAGAGGAGTGGCTCTCGATTCGCGACTATATCCATGTGATGACT 773
Db 838 CGTCGACTACCAACAGGAGCGGACCGGATCAGGATTAATACATGATGATGAGCT 897
QY 774 GGCAGATGGCATATGCTGCTCGCCCTCGAAAGCTCTTCAACGAGAGACATAGTTGTAC 833
Db 898 TGCCGATGGCCATTTGCCCACTGGAGAGCTCTTCGCTACTCTCTGACATTTGTTGT 957
QY 834 TGCTTACAACTGGGAACCTGGTCTGGAACATCTGTGCTTGAATGGTTACAGCATTTGA 893
Db 958 GGTATCAATCTAGGAACAGGTTGTGGAACAGGTTCTCGAGTGGTGAAGCGTTTGA 1017
QY 894 AAGGCTTCTGCAAGAAATTCAGTAAATTTATGTCAGAGACCGGAGATGCGAC 953
Db 1018 GGAGGGTTCGGAAGAAATTCATCAAGATTTGCCCCAGAGACCTGGAGATTGAC 1077
QY 954 TGAGGTTTATGATCTACAGAGAGCTGAGAAAGAACTTGTGGAAGGCAAACTATGG 1013
Db 1078 TGAGGTTTACCTTCACTGACAGGCGCAGAGAGGAGCTCGATGAGTGTCTCGTTGG 1137
QY 1014 TGTGAGGAGATGTCAGGAGCAATGGAATGGGCAAGAACTATCCCTGGGTTACGC 1073
Db 1138 AATAGAGGACATGTCAGGAGCACTGAGTGAATGGGCAAGAGAAATCCCTAGGATACAG 1197
QY 1074 GGGAGAGCT 1083
Db 1198 CGCAATGCT 1207

RESULT 14
AB097460
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) OsUGE-3 mRNA for putative
UDP-glucose 4-epimerase, complete cds.
ACCESSION
AB097460
VERSION
AB097460.1 GI:26106013
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1223
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/tissue_type="immature seed"
/country="Japan: Kyoto, Shimogamo"
1..1223
/gene="OsUGE-3"
65..1147
/gene="OsUGE-3"
/codon_start=1
/product="putative UDP-glucose 4-epimerase"
/protein_id="BAC41499.1"
/db_xref="GI:26106014"
/translation="MVGNGEGRKAGASVLVTGGAGVIGTHVLRLEKGFATVVVDN
FHSVPAIDRVRLIAGALSARLDIAGLKSDDMEKVFPAKRYDAVHIFAGLKV
GSEVAPQMYENNVAQTMNLSAMTKYKIKVIFSSATGATQPKETPCVEDSKLSA
LNFYGTTLVLENYFRQQAADPEMVIILRRYDPIGAHSGDIDGDPGPIPNLLPY
IQVAVGRPELVNYGVDPTRDGTARDYIHVVDLADGHLAALKLFPATPDTCVAY
NLGTGCTTVLEVVKAPEEASGKKIPKICPRPBGDTCEVYASTDKAKELGWSAREG
IEDCRDQNWAKNPFVGSANAEQN"
ORIGIN
Query Match 36.6%; Score 479.6; DB 8; Length 1223;
Best Local Similarity 66.6%; Pred. No. 1.4e-105;
Matches 686; Conservative 0; Mismatches 344; Indels 0; Gaps 0;
QY 54 CATTCTGTGTCACCGGTGGTGGCGTTTCATTTCGGACCCACACCGTGGTCTCTCAA 113
Db 106 CGTCTGTGTGACGGGGGAGCGGGGTACATCGGACGACACCGTGTCTCGGCTCTGGA 165
QY 114 AGCTGGCTTCAGCGTTTCAATATTCGACAATTCGATACTCCGTCATGGAAGCAGTGA 173
Db 166 GAAGGGTTCGCGGTTCACCGTCTCGACAACTTCCCACTCCGTCGCCGAGGCGCTCGA 225
QY 174 CCGGTCGCCCAAGTGGTGGCGCTCTGCTTCAGAACCTCCAAATTCACCCAGGGCGA 233
Db 226 CCGGTCGCCCTCATTCGCGGCGCGCGCTCTCGCCCGCTCTGACTTCATCGCCGGGA 285
QY 234 TCTCCGGAATAGGATGACTTCGAGAACTCTTCTCCAAACAACTTGTGATGCGGTAT 293
Db 286 TCTCAAGAGCAAGGACGACATGAGAGGTTGTCGCGCCCAAGAGGATGACGCGGTAT 345
QY 294 CCATTGCTGGCTTGAAGCGTTGCTGAAGGTTGCGAAGCGTTCGCTATTTTGA 353

Db 346 CCACCTTCGCGCGCTCAAGCGCTGGGGAGAGCGTCGCGCACCGCAGATGTACTACA 405
 Qy 354 TTTTAAATTTGGTTGGACCACTCACTACGAGTTTATGGCAAGTATAATTCAGAAA 413
 Db 406 GAACAACTGTCGGGACCACTGAACTCTATCTCCGCAATGACCAAGTACGCTGCAAGAA 465
 Qy 414 GATGGTTTCTCATCTATGCAACCGTTTATGCGCAACTGAAAGATACCGTGTGAGGA 473
 Db 466 GATAGTGTCTCTGTCGCGACGCGCTACGCGCAGCGGAGAGACCCCTCTGCTCGA 525
 Qy 474 GGATTTCAAGTTTACAAGCTATCAATCCCTATGCGGACCAAGCTTTCTTGGAGAAAT 533
 Db 526 GGATTTCAAGCTAGCGCTCTCAACCTATACGCGACCAAGCTCTGCTTGGAGAACTA 585
 Qy 534 TCCCGAGATATTCAGAAAGCTGAACCAAGATGGAAGATCATATPATCTGAGATCTTCAA 593
 Db 586 CTTCCGGCAGGTGCAGCGCGCCGACCCCGAGATGAGGCTGATCTCTGCTCAGTACTTCA 645
 Qy 594 TCCAGTTGGGCTCATGAAGTGGCAACTCGGTGAAGATCCCAAGGCGATCCCAATAA 653
 Db 646 CCCCATCGCGCTTCCCGGAGCGGCGACATCGGGAGGACCCCGGGGCTCCCAACAA 705
 Qy 654 CTTCTATGCTTCACTTACAGCAAGTATGTTGGAAAGATTGACTGAACCTCAATGTATACGG 713
 Db 706 CTTCTTCTGCTATCCAGCAGCTGCGCTGCGCGCCGCGCGAGCTCAAGCTCTACGG 765
 Qy 714 TCATGATTTACAAAGAGGATGGCTCTGCGATCCGGGACTATATCCATGTATGAGCTT 773
 Db 766 CTTGAGTACTACCAACAGGAGCGGACCGCGATCAGGGAATTACATACATGATGAGCT 825
 Qy 774 GGCAGATGCGCATTTGCTGCGCTCGAAAGCTTTTCAACAGGAGACATAGGTTGTAC 833
 Db 826 TGCGATGCGCACTTTCGCGCACTGGAGAGCTCTTTCGCTACTCTGCTGATGTTGTGT 885
 Qy 834 TGCTTACAACTGGGAACTGGTCTGGAAACATCTGTGTGTTGAAATGGTTACAGCATTTGA 893
 Db 886 GGCCTTACAACTAGGAACAGGCTGTGGAAACAGCTGCTCGAGTGGTGAAGGCGTTTCA 945
 Qy 894 AAAGGCTTCTGGCAAGAAATTCAGTAAATATGTTCCAGAAAGACCGGAGATGCGAC 953
 Db 946 GGGGGCTCGGAAAGAAATTCCTATCAGATTTGCCCGCAAGACCTGGAGATTGAC 1005
 Qy 954 TGAGTTTATGATCTACAGAGAGCTGAGAAAGAACTGTTGGAAAGGCAAACTATGG 1013
 Db 1006 TGAGTTTATGCTTCCACTGACAAAGGCAAGGAGGCTCGGATGGAGTGTGCTGCTTGG 1065
 Qy 1014 TGTGGAGAGATGTGCGAGGACCAATGGAATGGGCAAGACAACTCCCTGGGTTACGC 1073
 Db 1066 AATAGAGGACATGTGCGAGGACCAAGTGAATTTGGCCAAAGAAATCCGTACGGATACAG 1125
 Qy 1074 GGGGAAGCCT 1083
 Db 1126 CGCAATGCT 1135

RESULT 15
 BT009044
 LOCUS
 DEFINITION
 Triticum aestivum clone wdk5c.pk006.04:fls, full insert mRNA
 1516 bp mRNA linear PLN 20-JUN-2003
 sequence.
 ACCESSION
 BT009044
 VERSION
 BT009044.1 GI:32128595
 KEYWORDS
 FLI-CDNA.
 SOURCE
 Triticum aestivum (bread wheat)
 ORGANISM
 Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 1516)
 Tingey, S.V., Wolters, P., Powell, W., Dolan, M., Miao, G.-H.,
 Caraher, N.R., Hanafey, M.K. and Hainey, C.F.
 Direct Submission
 Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and

Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
 USA
 FEATURES
 source
 Location/Qualifiers
 1. .1516
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="wdk5c.pk006.04:fls"
 ORIGIN
 Query Match 36.3%; Score 476.4; DB 8; Length 1516;
 Best Local Similarity 66.4%; Pred. No. 8.8e-105;
 Matches 684; Conservative 0; Mismatches 346; Indels 0; Gaps 0;
 Qy 54 CATTCGCTTACCGGTGGTGGCGGTTTCAATGACCAATTCGATAATTCGATCACTCGTCAAGACAGTGA 113
 Db 12 CGTGTGTTGACCGCGCGCGGTTTCAATGCGCACGACACCGTCTGCGAGCTGTGGA 71
 Qy 114 AGCTGGCTTCAGCGTTTCAATAATGACAAATTCGATAATTCGATCACTCGTCAAGACAGTGA 173
 Db 72 GAAGGGCTAGCGCTCACCGCTCGACAACTTCGAACTCGTCCCGAGCGCTCGA 131
 Qy 174 CCGCTGCGCAAGTGGTGGCCCTCTGCTTCTCAGAACTTCAAGTCACTCAATTCACCCAGGCGGA 233
 Db 132 CCGCTGCGCACATCGTGGCGCCCGCTCTCGCGCGCTTCAATTCATCTTCGCGGA 191
 Qy 234 TCTCCGGAATAGGATGACTTGGAGAACTTCTTCTCCAAACAACTTTGATGCGCTGAT 293
 Db 192 CTTGACGATCAAGATGACTTGGAGAGTCTTTCGCGCCGCAAGAGTACGACCGCTGAT 251
 Qy 294 CCACCTTGTGCTTGAAGCGGTTCGAAAGCGTTCGAAAGCGTTCGAAAGCGCCGCTGCTATTGGA 353
 Db 252 ACATTCGCGCGCTCAAGCGGTGGCGGAGAGCTGGCGCACCCCGAGATGTACAACCG 311
 Qy 354 TTTTAATTTGGTTGGACCACTCAACCTCTACGAGTTTATGGCAAGTATAATTCGAAAA 413
 Db 312 CAACAACTCGTGGCACCGTCAACCTCTACGAGCTCATGAAGAGCAAGGTTGCAACAA 371
 Qy 414 GATGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTTGAAGAGTACCGTGTGAGGA 473
 Db 372 GTTGTGTTCTCGTCTCGCGACCGTGTACGCGCAGCGGAGAGTGCCTGCTTCA 431
 Qy 474 GATTTCAAGTTACAGCTATGAACTCCCTATGACGCGACCAAGCTTTCTCTGCAAGAAAT 533
 Db 432 GGAATCCCGCTCAAGCGCTTCAACCGTACCGCAGGACCAAGCTGTACCTGGAGGAT 491
 Qy 534 TGCCCGAGATATTCAGAAAGCTGAACCAAGATGGAAGATCATATTTACTGAGATCTTCAA 593
 Db 492 GCTGCGGACTACGAGCACGCGAACCGGAGTGGAGACGATCCTCTGCTGCTACTTCAA 551
 Qy 594 TCCAGTTGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCCAAGGCAATCCCAATAA 653
 Db 552 CCCCATCGCGCACACGAGAGCGCGACATCGGGGAGGACCCCAAGGGCGTCCCAACAA 611
 Qy 654 CCTCATGCTTACATTCAGCAAGTGTGTTGGAAGTGTGACTGAATCTCAATGTATACGG 713
 Db 612 CTTGCTCCCTCATCATCCAGAGTGTGCGCTGCGCCCGCGCCGAGCTCAACGCTTACGG 671
 Qy 714 TCATGATTTATCCAAAGAGGATGCTCTGCTGATCCCGGACTATATCCATGTGATGACTT 773
 Db 672 CCAGACTACCGCACCCGCGACCGGACCGCGCTCAGGGACTCATCCACGCTGCTGACCT 731
 Qy 774 GGCAGATGGCCATATTTGCTGCCCTCGGAAGCTTTTCAACAGGAGAACTAGGTTGTAC 833
 Db 732 CGCGAGCGCCCATCTCGCGCTTCAGAAAGCTTTTCGCCACCCCTGACATCGGCTGTGT 791
 Qy 834 TGCTTACAACTGGAACTGGTCTGGAAACATCTGTCTTGAATGTTTACAGATTTGA 893
 Db 792 GGGGTACAACCTGGGACCGGGCGGAGACGAGCTGTGGAGTGTGGAGTGTGAGCGCTTGA 851
 Qy 894 AAGGCTTCTGGCAAGAAATTCAGTAAATTTATGTCAGAGAGACCGGAGATCGGAC 953
 Db 852 GAAGGCATACGGCAAGAAATTCGCGTGAAGATGTGCCCGAGGAGCGCGGATTCGA 911

Qy	954	TGAGGTTTATGCACTACAGAGAGAGCTGAGAAAGAACTTGGTTGGAGGGCAAACTATGG	1013
Db	912	GCAGGTGTACGCGTCCACCGCCCAAGCCGAGAGAGGAGCTCGGCTGGAGGGCCCAAGTACGG	971
Qy	1014	TGTGGAGGAGATGTGACAGGACCAATGGAATTGGGCAAGAACTCCCTGGGGTTACGC	1073
Db	972	GATCGAGGAGATGTGACAGGACCAAGTGGAACTGGGCCAAGAGACCCCGTATGGCTACTG	1031
Qy	1074	GGGAAAGCCT	1083
Db	1032	CGGCAACGCT	1041

Search completed: March 22, 2004, 19:14:59
Job time : 5330 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 17:28:33 ; Search time 572 Seconds
(without alignment)
9744.119 Million cell updates/sec

Title: US-09-913-064a-13
Perfect score: 1312
Sequence: 1 gcacagcactctctctcc.....aaaaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Watch 0%
Maximum Watch 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1312	100.0	1312	3	AAA59439 Nucleotid
2	644.4	49.1	1056	6	Abz14552 Arabidops
3	629.4	48.0	1397	3	AAC34854 Arabidops
4	603.2	46.0	1488	6	ABA00335 Potato ps
5	601.8	45.9	1056	6	ABA00336 Potato ps
6	544.2	41.5	1508	3	AAC37107 Arabidops
7	476.4	36.3	1516	3	AAA59440 Nucleotid
8	445	33.9	1065	7	ADA69621 Rice Gene
9	445	33.9	1498	3	AAA59442 Nucleotid
10	431.4	32.9	1532	3	AAA59443 Nucleotid
11	431.2	32.9	1394	3	AAC51240 Arabidops
12	431.2	32.9	1396	3	AAC39471 Arabidops
13	428.4	32.7	1286	3	AAC36878 Arabidops
14	424.8	32.4	479	3	AAA59435 Nucleotid
15	423.6	32.3	1292	3	AAC47345 Arabidops
16	416	31.7	1314	3	AAC50699 Arabidops
17	412	31.4	1314	3	AAC47111 Arabidops
18	393.4	30.0	1393	3	AAA59441 Nucleotid
19	385.6	29.4	1335	7	ADA70308 Rice Gene
20	312	23.8	641	6	ABQ66038 Arabidops
21	278.2	21.2	4626	2	AAX84321 Stealth v
22	277.8	21.2	1017	6	ABN89338 B. subtil
23	277.8	21.2	5793	6	AA517998 Plasmid p

24	277.8	21.2	5943	6	AA517997
25	267.6	20.4	1023	7	ACP70498
26	267.6	20.4	69727	7	ACF65374
27	267.6	20.4	110000	7	ACF67367_35
28	262.4	20.0	1008	7	ABX07603
29	262.4	20.0	110000	7	ABX07603 S. pneumo
30	262.2	20.0	1050	8	ABX09547
31	262.2	20.0	110000	8	ADB09547
32	261.8	20.0	100848	4	ADB12064_09
33	261.6	19.9	1017	7	AA728552_09
34	261.6	19.9	13206	2	AAV52166
35	261.6	19.9	110000	7	AAV52166 Streptoco
36	261.6	19.9	110000	7	ABX07339
37	248.6	18.9	545	3	AAV52166
38	247.2	18.8	3140	2	AAV52291
39	244	18.6	1047	3	AAZ45319
40	244	18.6	1488	6	ABL62872
41	244	18.6	1488	6	ABL67513
42	244	18.6	1488	6	ABL63090
43	244	18.6	1488	6	ABL67063
44	241.4	18.4	110000	2	AA742063_03
45	240.8	18.4	1249	2	AAT58301

ALIGNMENTS

RESULT 1
AAA59439
ID AAA59439 standard; DNA; 1312 BP.

AC AAA59439;

DT 14-NOV-2000 (first entry)

DE Nucleotide sequence of UDP-galactose 4-epimerase.

XX UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
KW raffinoseaccharide pathway; raffinose; stachyose; carbohydrate metabolism;
KW starch level; ss.

OS Glycine max.

PH Key Location/Qualifiers
FT CDS 415..1086

FT /tag= a
FT /product= "UDP-galactose 4-epimerase"
FT /note= "the codons encoding amino acids 3-129 are not given"

XX WO200047755-A2.

XX PD 17-AUG-2000.

XX PF 09-FEB-2000; 2000WO-US003453.

XX PR 10-FEB-1999; 99US-0119588P.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Carlson TU, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;

XX PI Rafalski JA, Thorpe CJ;

XX DR WPI; 2000-549152/50.

XX DR P-PSDB; AAB07819.

XX PT Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant cells to e.g. decrease cell wall constituents and increase starch content.

XX PS Claim 2; Page 51; 61pp; English.

XX CC The present sequence encodes a UDP-galactose 4-epimerase. The enzyme is

CC responsible for the interconversion of UDP-glucose and UDP-galactose. As
 CC the activity of the enzyme appears to be particularly limiting to carbon
 CC flux into the raffinose/saccharide pathway, reduction of its activity should
 CC decrease the levels of raffinose and stachyose in seeds. The
 CC polynucleotide is useful to transform plant cells to enhance the degree
 CC of expression of UDP-galactose-4-epimerase within these cells. Changes in
 CC the expression of UDP-galactose-4-epimerase within a cell enables
 CC modification of plant carbohydrate metabolism. This allows growth of
 CC grains with reduced cell wall constituents (fiber) and increased levels
 CC of starch
 XX

SQ Sequence 1312 BP; 386 A; 273 C; 300 G; 353 T; 0 U; 0 Other;

Query Match 100.0%; Score 1312; DB 3; Length 1312;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGCCACTTCTCCCTCTCTATTGCGACATGGTGTCTCTCCCAACACATTTCTG 60
 DB 1 GCACGAGCCACTTCTCCCTCTCTATTGCGACATGGTGTCTCTCCCAACACATTTCTG 60
 QY 61 GTCACGGTGTGCGGTTTCATTGGCAACCCACACCGTCTGATCTCTCAAGCTGGC 120
 DB 61 GTCACGGTGTGCGGTTTCATTGGCAACCCACACCGTCTGATCTCTCAAGCTGGC 120
 QY 121 TTCAGCGTTTCAATATCGACAATTCGATACCTCCGTCATGGAGCAGTGGACCGGTC 180
 DB 121 TTCAGCGTTTCAATATCGACAATTCGATACCTCCGTCATGGAGCAGTGGACCGGTC 180
 QY 181 CGCCAAAGTGGTGGCCCTCTGTTTCTCAGAACCTTCCAAATTCACCGGGCGATCTCCGG 240
 DB 181 CGCCAAAGTGGTGGCCCTCTGTTTCTCAGAACCTTCCAAATTCACCGGGCGATCTCCGG 240
 QY 241 ATATGGGATGACTTGGAGAACTCTTCTCCAAACCAATTTGATGCGCGTGCATCCACTTT 300
 DB 241 ATATGGGATGACTTGGAGAACTCTTCTCCAAACCAATTTGATGCGCGTGCATCCACTTT 300
 QY 301 GCTGGCTTGAAGCGGTTGCTGAAAGCGTTGCGAAGCCCGTCCGCTATTTTGAATTAAT 360
 DB 301 GCTGGCTTGAAGCGGTTGCTGAAAGCGTTGCGAAGCCCGTCCGCTATTTTGAATTAAT 360
 QY 361 TTGTTGGACCATCACTCTACAGTTTATGGCAAGATATATGGCAAGATGGTT 420
 DB 361 TTGTTGGACCATCACTCTACAGTTTATGGCAAGATATATGGCAAGATGGTT 420
 QY 421 TTCTCATCTCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGGAGATTC 480
 DB 421 TTCTCATCTCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGGAGATTC 480
 QY 481 AAGTTACAAGCTATGAATCCCTATGGACGCAAGCTTTCTGGAAGAAATTCGCCGA 540
 DB 481 AAGTTACAAGCTATGAATCCCTATGGACGCAAGCTTTCTGGAAGAAATTCGCCGA 540
 QY 541 GATATTCAGAAAGCTGAACAGATGGAAGATCATATTACTGAGATCTTCAATCCAGTT 600
 DB 541 GATATTCAGAAAGCTGAACAGATGGAAGATCATATTACTGAGATCTTCAATCCAGTT 600
 QY 601 GGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAATTAACCTCATG 660
 DB 601 GGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAATTAACCTCATG 660
 QY 661 CCTTACATTCAGCAAGTACTGCTGGAGATGACTGCACTCAATGATATCGGTCAATGAT 720
 DB 661 CCTTACATTCAGCAAGTACTGCTGGAGATGACTGCACTCAATGATATCGGTCAATGAT 720
 QY 721 TATCCAAAGGAGGATGGCTCTCGCATCCGGGACTATATCCATGTGATGGACTTGGCAGAT 780
 DB 721 TATCCAAAGGAGGATGGCTCTCGCATCCGGGACTATATCCATGTGATGGACTTGGCAGAT 780
 QY 781 GGCCATATTCGCTCCCTGGAAAGCTTTCACAAAGGAGATAGGTGACTGCTTAC 840
 DB 781 GGCCATATTCGCTCCCTGGAAAGCTTTCACAAAGGAGATAGGTGACTGCTTAC 840

QY 841 AACCTGGGAACGTGTCGTGGAACATCTGTGTTGAAATGGTTACAGCATTTGAAAGGCT 900
 DB 841 AACCTGGGAACGTGTCGTGGAACATCTGTGTTGAAATGGTTACAGCATTTGAAAGGCT 900
 QY 901 TCTGGCAAGAAATTCACGATTAATATGTCCAAAGACCGGAGATGCGACTGAGGTT 960
 DB 901 TCTGGCAAGAAATTCACGATTAATATGTCCAAAGACCGGAGATGCGACTGAGGTT 960
 QY 961 TATGCACTACAGAGAGCTCGAGAAAGAACTTGTGTGGAAGGCAAACTATCGTGTGGAG 1020
 DB 961 TATGCACTACAGAGAGCTCGAGAAAGAACTTGTGTGGAAGGCAAACTATCGTGTGGAG 1020
 QY 1021 GAGATGTCAGGAGCCCAATGGAATTTGGCAAGAACATCCCTGGGTTTACCGGGGAAG 1080
 DB 1021 GAGATGTCAGGAGCCCAATGGAATTTGGCAAGAACATCCCTGGGTTTACCGGGGAAG 1080
 QY 1081 CCTTGAATTTAGCTTTGAGAAATATATCTGCTCATCTAGCAATGCTTTTCAATATAAGCA 1140
 DB 1081 CCTTGAATTTAGCTTTGAGAAATATATCTGCTCATCTAGCAATGCTTTTCAATATAAGCA 1140
 QY 1141 TCTCTTATATAGATTAATTTTATGTTTGTATGATTTTGTAGGAGTTCGTTGTATATCT 1200
 DB 1141 TCTCTTATATAGATTAATTTTATGTTTGTATGATTTTGTAGGAGTTCGTTGTATATCT 1200
 QY 1201 TGACAATAAAAAATTTGGCAGCATTTTCAAGAAAGTAAAGCTATGTATTTAAACAATAACT 1260
 DB 1201 TGACAATAAAAAATTTGGCAGCATTTTCAAGAAAGTAAAGCTATGTATTTAAACAATAACT 1260
 QY 1261 TAAATTAGACTGGCCATTTGATTTGATTTGAAAAAATAAAAAA 1312
 DB 1261 TAAATTAGACTGGCCATTTGATTTGATTTGAAAAAATAAAAAA 1312

RESULT 2

ABZ14552 standard; DNA; 1056 BP.

XX ABZ14552;

XX 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 2357.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

XX 26-JAN-2001; 2001US-0264647P.

XX 22-JUN-2001; 2001US-0300111P.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and
 XX producing plants with increased tolerance to these abiotic stresses.
 XX Claim 144; SEQ ID NO 2357; 577pp + Sequence Listing; English.
 XX The invention relates to identifying a stress condition to which a plant
 XX cell has been exposed, comprising: (a) contacting nucleic acid
 XX representative of expressed polynucleotides in the plant cell with an
 XX array or probes representative of the plant cell genome; and (b)

PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140153P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 26-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-014287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-014320P.
PR 12-JUL-1999; 99US-0143277P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160860P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match

48.0%; Score 629.4; DB 3; Length 1397;

[illegible]

Db 282 TTGATAGATCAGAGAAATTAGTTGGTCTCAACTTTTCAGAGAACTTTCGAAATTCGATTTGG 341
 QY 230 GCGATCTCCGGAATAGGATGACTTGGAGAACTCTCTCCAAACACACATTTGATGCCG 289
 Db 342 GTGATATTAGAAACAAGATGCTTGGAGAGCTATTTCTTAAGAAAGAGTTTGTCTGTG 401
 QY 290 TGATCCACTTGTCTGGCTTGAAGGGTGTCTGAAAGGGTTTCGGAAGCCCGTCTGCTATT 349
 Db 402 TGGTCCATTTTGTGACTTAAGCTGTGGAGAGAGTGTGTTCAGGCCCTTCTTTACT 461
 QY 350 TTGATTTTAATTTGGTGGCCACCATCAACCTCTAGAGTTTATGCGAAGATATATTCGA 409
 Db 462 TTGAGAACCAATCTGATTGGATCAATAACTTTGTATTTCAGTCTATGCCCAAGTATATTTGA 521
 QY 410 AAAAGATGTTTCTCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTG 469
 Db 522 AGAGTTGGTTTTTTCATCATCTGCGACAGTTTATGGTCAGCTGAAAGGTTCCCTGTG 581
 QY 470 AGGAGATTTCAGTTTCAAGCTATGAATCCCTATGGACGGAACCAAGCTTTCTGGGAAG 529
 Db 582 TGGAGGATTTTGAATTTGAAGGCTATGAATCCTTATGGTCGAACAAAGCTATTTCTTGAAG 641
 QY 530 AATATCCCGGAGATATTCAGAAAGCTGAACCAAGATGGAAGATCATATTACTGAGATACT 589
 Db 642 ATATTCCTCGGATATCCAGAGGAGATCAAGATGGATATCATACTGTTGAGGTATT 701
 QY 590 TCAATCCAGTTGGGCTCATGAAAGTGGCAAACTCGTGAAGATCCCAAGGGCATCCCA 649
 Db 702 TCAACCCAGTAGGAGCTCATGAAAGTGGCAAACTCGGGAAGATCCAAAGGGCATTCGA 761
 QY 650 ATACTCTAGCTTACATTCAGCAAGTACGTTGGAGAGTTGACTGAACCTCAATGAT 709
 Db 762 ACAATCTTAGCTTACATTCAGCAAGTACGTTGGTAGATTCGAGAGTTGAATGAT 821
 QY 710 ACGGTCTATGATTTCAACAGAGGAGTGGCTCTGCGATCCGGAATATATPCCATGTGATGG 769
 Db 822 ATGGCAAGCACTACCTACACCTGATGGTACCGGATACGAGATTATATCCATGTTTGG 881
 QY 770 ACTTGGCAGATGCCATATGCTGCCCTCGGAAGCTTTCAACGGAGACATAGTT 829
 Db 882 ATTAGCGGACGCTCATGTTGTTGCACTTCAGAGACTTCTTAGGCGAATCATTTAGGTT 941
 QY 830 GTACTGTGTACAACCTGGGAACCTGGTCTGTGGAACATCTCTGTCTGAAATGGTTACAGCAT 889
 Db 942 GTGTGCTTCAATTTGGGTACTTGGAAAGGCAAACTCTCTAGAGATGGTTGCTGCTT 1001
 QY 890 TTGAAAGGCTTCTGGCAAGAAATTCAGTAATATATGTCGAAGAACCGGGAGATG 949
 Db 1002 TTGAAAGAGCGTCTGGAAGAAAAATCCGCTTAAATGTGTCCAGAACACCGAGAGATG 1061
 QY 950 CGACTGAGTTTATGCATCTACAGAGAGCTGAGAAAGAACTTGTGTGGAAAGGCAAACT 1009
 Db 1062 CCAGTCTGTTTATGCATCTACTGAAAGAGCTGAGAGAGCTCGGTTGGAAGGCAAAAT 1121
 QY 1010 ATGGTGGAGGAGATGTGAGGACCAATGGAAATTTGGGCAAGAAACATATCCCTGGGTT 1069
 Db 1122 ATGATATAATGATGATGTGAGGAGACAGTGGAAATGGGCAAGCAAAATCTCTGGGTT 1181
 QY 1070 ACGGGGGAAGCTTGAATTAGCTTGAAGAAATATAC 1105
 Db 1182 ACCAATCAAGCCCTTGAATGTTAAATTAATATAC 1217

RESULT 5

ABA00336/c

ID ABA00336 standard; cDNA; 1056 BP.

XX ABA00336;

AC ABA00336;

XX 10-DEC-2002 (first entry)

XX Potato psen-1 antisense sequence.

XX

Potato; psen-1; 5'-diphospho-galactose 4-epimerase; transgenic; DGE;
 uridine diphosphate glucose epimerase; UDPG; carbohydrate metabolism;
 Controlled Environment Agriculture; nutritional profile; antisense; ss.
 Solanum tuberosum.

WC200270649-A2.

12-SEP-2002.

01-FEB-2002; 2002WO-US002727.

01-FEB-2001; 2001US-0265311P.

(DAIZ/) DAI Z.

Dai Z, Shi L, Hooker BS;

WPI; 2002-698747/75.

New isolated polynucleotide molecule comprising a gene coding for the

enzyme 5'-diphospho galactose 4-epimerase, useful for regulating

carbohydrate metabolism or nutritional profile in transgenic plants.

Example 5; Fig 3; 38pp; English.

This sequence represents the antisense cDNA sequence of potato psen-1

which encodes the enzyme 5'-diphospho-galactose 4-epimerase (DGE). The

DGE sequence is homologous to the sequence of A. thaliana uridine

diphosphate glucose (UDPG) epimerase gene (86% homology at peptide level,

71% homology at the nucleotide level). The DGE polynucleotide molecules

and antisense sequences are useful in regulating carbohydrate metabolism

in transgenic plants. The polynucleotide can be used in controlled

Environment Agriculture that employs an integrated system for commercial

production of transgenic plants in a controlled environment. Controlling

the level of DGE in transgenic plants is useful for regulating the

nutritional profile of the plant

Sequence 1056 BP; 311 A; 263 C; 174 G; 308 T; 0 U; 0 Other;

Query Match 45.9%; Score 601.8; DB 6; Length 1056;

Best Local Similarity 73.8%; Pred. No. 1.3e-144;

Matches 765; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

QY 50 AACACATTTCTGGTCCCGTGGTCCCGTGGTTCATTTGTCACCCACACCGTGTTCAGCTTC 109

Db 1037 AAAATATTTTGGTACTGCGAGGAGCTGTTTTCATTTGGAACACACACTGTGTGCGATTAC 978

QY 110 TCAAGCTGCTTCAGCGTTTCAATATCGCAATTTTCGATACTCCGTCATGGAAGCAG 169

Db 977 TGAATGAAGGTTTCAAGTTTACCATTATGATTAATTTTCAATTTCTGTGAGAGACTG 918

QY 170 TGGACCGCTCCGCAAGTGGTTGGCCCTCTGCTTTCTCAGAACCTCCAAATTCACCCAGG 229

Db 917 TTGATAGATCAGAGAATTAGTTGGTCTCAACITTTACAGAAATCTTGAATTCATTGG 858

QY 230 GCGATCTCCGAATAGGATGACTTGGAGAACTCTTTCBAACACACATTTTCATGCCG 289

Db 857 GTGATATTAGAAACAAGATGACTTGGAGAGCTATTTTCTAAGAAAGAGTTTCTGCTG 798

QY 290 TGATCCACTTGTCTGGCTTGAAGCGGTGTGTAAGCGTTGTGGAAGCCCGTCCGCTATT 349

Db 797 TGGTCCATTTTGTCTGGACTTAAAGCTGTGTGAGAGAGTGTGTTCAGCCCTTTCTTACT 738

QY 350 TTGATTTTAATTTGGTGGCCACCATCAACCTCTACGAGTTTATGGCAAGATATATTCGA 409

Db 737 TTGAGAACCAATCTGATTGGATCAATACTTTGTATTTCAGTCATGGCCCAAGTATATTTGA 678

QY 410 AAAAGATGTTTCTCTCATCTGCAACCGTTTATGGCCAACTTCGCAAGATACCGTGTG 469

Db 677 AGAAGTTGGTTTTTTCATCATCTGGACAGTTTATGTGTGTCAGCTTGAAGAGTTTCCCTGTG 618

QY 470 AGGAGATTTCAGATTACAGCTATGAATCCCTATGGACGGAACCAAGCTTTCTCTGGAAG 529

```
Db      617  TGGAGGATTTGAATTTGAAGGCTATGAATCTTATGGTCGAACAAAGCTATTTCCTTGAAG 558
Qy      530  AATTTCCCGAGATATTCAGAAAGCTGACACCGAATGGAGATCATATTTACTGAGATCT 589
Db      557  ATATTCTCGGATATCCAGAGGCGAGATCAAGATGGAATATCACTGTTGAGGTATT 498
Qy      590  TCAATCCAGTTGGGGTCTATGAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAA 649
Db      497  TCAACCCAGTAGGAGCTCATGAAGTGGCAAACTCGGGAAGATCCAAAGGSCATCCCA 438
Qy      650  ATAACTCATGCTTACATTCAGCAAGTAGCTGTTGGAAAGTAGTGAAGTCAATGAT 709
Db      437  ACAATCTTATGCTTACATTCAGCAAGTAGCTGTTGGTAGATTGCCAGAGTTGAATGAT 378
Qy      710  ACGGTCTATGATATCAACGAGGAGTGGCTCTGCGATCCGGACTATATCCATGTGATGG 769
Db      377  ATGGCAACGACTACCTTACACCTGATGGTACCGGATACGAGATTATATCCATGTTTGG 318
Qy      770  ACTTGCAGATGGCCATATGTGCTGCGCTCGAAAGCTTTTCACACCGGAGAACATAGGTT 829
Db      317  ATTAGCGGACGCTCATGTTGTTGCACTTCAGAGACTTCTAAGGCGAGATCATTTAGGTT 258
Qy      830  GTACTGCTTACACCTCGGAAGTGTGCGGGAACATCTGTGCTTGAATGGTTACAGCAT 889
Db      257  GTGTTGCTTACAAATTTGGGTACTGGAAGGCAAACTCTGCTAGAGATGGTTGCTGTT 198
Qy      890  TTGAAAGGCTTCTGCAAGAAATATCCAGTAAATTTATGTCACAGAACCGGAGATG 949
Db      197  TTGAAAGGCTCTGGAAGAAATCCCGCTTAAATGTGTCGAAGAACCGAGAGATG 138
Qy      950  CGACTGAGTTTATGATCTACAGAGAGCTGAGAAAGAACTTGGTTGGAGGCAAACT 1009
Db      137  CCACTGCTGTTATGATCTACTGAAAAGCTGAGAAAGAGCTCGGTTGGAAAGGCAAAAT 78
Qy      1010  ATGGTGTGAGGAGATGTGAGGGACCAATGGAATGGGCAAGAACAAATCCCTGGGGTT 1069
Db      77  ATGGTATAATGAGATGTGAGGACCACTGGAATGGGCAAGAACAAATCCCTGGGGTT 18
Qy      1070  ACGCGGGAGCCTTGA 1086
Db      17  ACCAATCAAAGCCTTGA 1
RESULT 6
AAC37107
ID AAC37107 standard; DNA; 1508 BP.
XX AC AAC37107;
XX XX
XX 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 16200.
XX XX
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway; metabolic pathway;
XX KW promoter; termination sequence; ss.
XX XX
XX OS Arabidopsis thaliana.
XX PN EF1033405-A2.
XX XX
XX PD 06-SEP-2000.
XX XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX XX
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
```

```
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134376P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136352P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
```

[illegible]

Db 822 AATAACCTCATGCTTACATCCAAAGTGGCGTTGGAGCTTTACCGGAACCTCAATGTC 881
Qy 709 TAGGTCATGATATCCACGAGGATGGCTCTGGATCCGGACTATATCATCTGATG 768
Db 882 TAGGACATGATATCCACGAGGATGGTAGTGGCTAAGAGACTTACATCCATGTGATG 941
Qy 769 GACTTGGCAGATGGCCATATGTCGCCCTCGGAAGCTCTTCACACCGGAGAACATAGT 828
Db 942 GATTAGCAGATGGCCATATGCTGGCTCAGGAGCTATTTGCTGATCCAAAGATTGGT 1001
Qy 829 TGTACTGCTTACNACTGGGAAGTGGTCTGGACATCTGCTTGAATGTTACACA 888
Db 1002 TGTACTGCTTACNACTTAGGACTGGTCAAGGAGCTGTGTGTAAGATGTTGAGCT 1061
Qy 889 TTTGAAAAGCTTCTGGCAAGAAAATCCAGTAAATTTATGTCGAAGAGACCGGAGAT 948
Db 1062 TTTGAAAAGCTTCCGGCAAGAAAATCCGATTAGCTCTGTCCGAGAGGTTCAGGAT 1121
Qy 949 G 949
Db 1122 G 1122

RESULT 7
AA59440
ID AA59440 standard; DNA; 1516 BP.
XX
AC AA59440;
XX
DT 14-NOV-2000 (first entry)
XX
DE Nucleotide sequence of UDP-galactose 4-epimerase.
XX
KW UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
KW raffinose pathway; raffinose; stachyose; carbohydrate metabolism;
KW starch level; ss.
OS
OS Triticum aestivum.

XX Key Location/Qualifiers
FH CDS 1..1059
FT /*tag= a
FT /product= "UDP-galactose 4-epimerase"
XX WO200047755-A2.
PD 17-AUG-2000.
XX
XX 09-FEB-2000; 2000WO-US003453.
XX
XX 10-FEB-1999; 99US-0119588P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Carlson TJ, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;
XX Rafalski JA, Thorpe CJ;
XX WPI: 2000-549152/50.
XX P-PSDB; AAB07820.
XX
XX Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant
XX cells to e.g. decrease cell wall constituents and increase starch
XX content.
XX
XX Claim 2; Page 52; 61pp; English.
XX
XX The present sequence encodes a UDP-galactose 4-epimerase. The enzyme is
XX responsible for the interconversion of UDP-glucose and UDP-galactose. As
XX the activity of the enzyme appears to be particularly limiting to carbon
XX flux into the raffinose pathway, reduction of its activity should
XX decrease the levels of raffinose and stachyose in seeds. The
XX polynucleotide is useful to transform plant cells to enhance the degree

CC of expression of UDP-galactose-4-epimerase within these cells. Changes in
CC the expression of UDP-galactose-4-epimerase within a cell enables
CC modification of plant carbohydrate metabolism. This allows growth of
CC grains with reduced cell wall constituents (fiber) and increased levels
CC of starch

XX
SQ Sequence 1516 BP; 374 A; 439 C; 472 G; 229 T; 0 U; 2 Other;
Query Match 36.3%; Score 476.4; DB 3; Length 1516;
Best Local Similarity 66.4%; Pred. No. 2.9e-112;
Matches 684; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

Qy 54 CATCTCTGTCACCGTGGTCCGCTTTCATTGGCCACCACACCGTCTGTCAGCTTCTCAA 113
Db 12 CGTCTGTCACCGCGCGCGGGGTTTCGCGACGACACCGTCTGTCAGCTCTCGGA 71
Qy 114 AGCTGGCTTTCAGCGCTTTCATTAATTCGACAATTCGATAACTCCGTCTATGAAGCAGTGA 173
Db 72 GAAGGCTACCGCTGACCGCGCTCGAACAATTCACAACCTCCGTCGCCGAGGCGCTCGA 131
Qy 174 CCGCTCGGCAAGTGGTGGCCCTCTGCTTCTCAGAACTCCAAATTCACCCAGGCGGA 233
Db 132 CCGCTCGGCAATCTGTCGCGCCCGCTCTCGCGCCCTCCAAATTCATCTTTCGGGGA 191
Qy 234 TCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAAACAACATTTGATGCCGTGAT 293
Db 192 CTTGACATCAGATGACTCTGAGAGGTCTTCGCGCCCAAGAGTACGACGCGGTGAT 251
Qy 294 CCATTTCTGCTGTAAGCGGTGCTGAAAGCTTGAAGCGTTGCGAAGCCCGTCTATTTTGA 353
Db 252 ACCTTCCGCGGCTCAAGCGGTGGCGGAGAGCGTGGCGCACCCCGGAGATGTACAACCG 311
Qy 354 TTTTAATTTGTTGGCACCATCAACCTCTAGAGTTTATGCAAGTATATTTGCAAAA 413
Db 312 CAACAACATCTGCGACCGCTCAACCTCTACAGCTCATGAGNAGCACGGTGCACAA 371
Qy 414 GATGGTTTCTCATCTGCAACCGTTTATGGCCAACTTGAAGATATACCTGTGAGGA 473
Db 372 GTTGGTTTCTGCTGCTGCGGACCGTGTACGCGCAGCGGAGAGTGGCTCTCTCGA 431
Qy 474 GGATTTCAAGTTTACAAGCTATGATCCCTATGGAGCGGACCAAGCTTTTCTGGAAGAA 533
Db 432 GGATCTCCCCCTCAAGGCGCTCAACCGCTACGCGAGGACCAAGCTGTACCTGGAGGAGAT 491
Qy 534 TGCCCGAGATATTCAGAAAGCTGAACAGAAATGGAAGATCATATTAATGAGATATCTCAA 593
Db 492 GCTGCGGAGTACCAGCAGCGAACCAGGAGTGGAGGAGCATCTGCTGGCTACTTCAA 551
Qy 594 TCCAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAGAGTCCCAAGGCGATCCCAATAA 653
Db 552 CCCATCGGCGCACACGAGAGCGCGGACATCGGGGAGGACCCCAAGGCGCTCCCAACA 611
Qy 654 CCTCATGCTTTACATTCAGCAAGTAGCTGTTGGAAGATTGACTGAACCTCAATGTATACGG 713
Db 612 CTTGCTCCCTTACATCCAGCAGGTGGCGCTCGCCCGCGCGGCTCAACGTCTACGG 671
Qy 714 TCATGATATCCAAAGGAGTGGCTCTGCAATCCGGAGCTATATCCATGTGATGGACTT 773
Db 672 CCAGACTTACCGCACCCGCGAGCGCACCGCGCTCAGGAGTACATCCAGTGGTGCACCT 731
Qy 774 GGCAGATGGCCATATTTGCTGCCCTGCGAAAGCTCTTCACAACGAGAACTAGGTGTGTAC 833
Db 732 CGCCGAGCGCCATCATCGGGGCTCGAGAGCTCTTCGCCACCCCTGACATCGGCTGTGT 791
Qy 834 TGTATTACAACTGGGAACTGCTCTGGAACAATCTGTGCTTGAATGTTTACAGCATTTGA 893
Db 792 GCGGTACAACTGGGAGCGGGGCGCGGAGCGACCGTGTCTGGAGATGGTGGAGCGGTTCGA 851
Qy 894 AAAGCTTCTGCAAGAAAATTTCCAGTAAATTTATGTCGAAGAGACCGGAGATGGCAC 953
Db 852 GAAGCATACGCGAAGAAAATTCGCGTGAAGATGTGCCCGAGGAGCGCGCGATTCGGA 911
Qy 954 TGAGTATTATCATCTACAGAGAGCTGAGAAAAGAACTTGGTTGGAAGGCAAACTATGG 1013

```

Db 912 GCAGGTGTACGGTCCACCGCCACAGGCGAAGAGAGCTCGGCTGGAGGGCCAAAGTACGG 971
QY 1014 TGTGAGGAGATGTCCAGGACCAATGAATTTGGCAAGAAACAATCCCTGGGTTACGC 1073
Db 972 CATCGAGGAGATGTCCAGGACCAAGTGAATCTGGGCCAAGAAGAACCGGTATGGCTACTG 1031
QY 1074 GGGGAAGCCT 1083
Db 1032 CGGCAACGCT 1041

RESULT 8
ADA69621
ID ADA69621 standard; DNA; 1065 BP.
XX
AC ADA69621;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 2944.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW Gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
XX WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 2944; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 1065 BP; 257 A; 298 C; 289 G; 221 T; 0 U; 0 Other;

Query Match 33.9%; Score 445; DB 7; Length 1065;
Best Local Similarity 66.1%; Pred. No. 3.1e-104;
Matches 675; Conservative 0; Mismatches 340; Indels 6; Gaps 2;

QY 55 ATTCGGTCCAGGTTTCAATATCGACAAATTCGATAACTCGTTCATGCAAGCAGTGGAC 114
Db 25 ATTCGGTCCAGGTTTCAATATCGACAAATTCGATAACTCGTTCATGCAAGCAGTGGAC 84
QY 115 GCTGGCTTCAGGTTTCAATATCGACAAATTCGATAACTCGTTCATGCAAGCAGTGGAC 174
Db 85 CTCGGCTTCGGGTTGTGCTCTCGACAACTCGACACGCTCCGAGCTCGCCATCTCTC 144

```

```

QY 175 CGGTCGCGCAAGTGGTTGGCCCTCTGCTTTCTCAGAACCTCCAATTACCCAGGGCGAT 234
Db 145 CGGTCAGGGAACCTGCCCGG---ACACAAAGCCAAACACCTCGACTTCGCAAGGTGAC 201
QY 235 CTCGGAATAGGAGATGACTTTGGAGAACTCTTTCTCCAAAACAACTTTGATCCCGTATC 294
Db 202 CTCGCGGACAAAGCAAGCGTTGGACCAATCTTCTCTCTCAAAGGTTTGAGGCTGTCTC 261
QY 295 CACTTTGCTGGCTTGAAGCGGTTGCTGAAGCGTTGGAAGCGGCTGCTATTTTGAT 354
Db 262 CATTTTGGCCGGCTGAAAGCTGTTGGCGAGAGCGGTGCAAGACCCCTGCTTACTACGAC 321
QY 355 TTTAATTTGGTTGGCAACCACTCAACCTCTACGAGTTTATGCGCAAGTATATAATTCGAAAAG 414
Db 322 AACAACTCATCGGCACCATCACTCTCTCTGAGGTATGCGCGCACATGGCTGCACCAAG 381
QY 415 ATGTTTTTCTCATCTCGAACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGGAG 474
Db 382 CTGGTGTCTCATCTCGCAACTGTCTACGGGTGGCCCAAGGAGGTGCCCTGCACTGAA 441
QY 475 GATTTTCAAGTTTCAAGCTATGAATCCTATGACGAGCAACCAAGCTTTTCTCGAAAGAAAT 534
Db 442 GAATCCCACTTTGTGCAATGAACCCCTACGCGAGACAAAGCTGGTATCGAGACATG 501
QY 535 GCGCGAGATATTGAGAAAGCTGAAACAGAAATGGAAGATCATATTACTGAGATCTTCAAT 594
Db 502 TGCGGGATCTGCATGCTTCAGACCCAAACTGGAAGATCATACTGCTCCGACTACTCAAC 561
QY 595 CAGTTTGGGGCTCATGAAAGTCGCAAACTCGGTGAAGATCCCAAGGGATCCCAATAAC 654
Db 562 CTGTTTGGAGCTCACCCAAAGCGGTACATTTGTTGAGACCCCTGCGGATCCCAACAC 621
QY 655 CTCATGCCCTTACATTCAGCAAGTACGCTGTTGAAAGATTGACTGAACTCAATGTATACGGT 714
Db 622 CTCATGCCCTTCTGTCAGCAGGTGCTGTTGGCAGAGGCGCGCTTACCCTCTATGGA 681
QY 715 CATGATTATCCACAGAGGATGCTCGATCCGGGACTATATCATGTGATGACTTG 774
Db 682 ACCGATCAACACCAAGGATGAACTGGGGTTCTGACTATATCATGTTGTTGATCTA 741
QY 775 GCAGATGGGCATATTGCTGCCCTCGGAAAGCTCTTCACA---ACGAGAACATAGGTTGT 831
Db 742 GCGGATGGTCAATATCGCCGCTTAAGAAAGCTCTATGAAGATCTCATAGAATAGGATGT 801
QY 832 ACTGCTTACAACCTGGAACTGCTGCGAACATCTGTGCTTGAATGTTTACAGCATTT 891
Db 802 GAGGTGTACAATCTGGGCACCTGAAAGGGGACATCTGTCTGGAATGTTGACGATTC 861
QY 892 GAAAAGGCTTCTGGCAAGAAAATTCAGTAAATATATGTCCAAAGAACCGGGAGATGCG 951
Db 862 GAGAAAGCTTCTGGAAGAAAATCCCGCTTGTATTGCTGGACGAAGGCTCGAGATGCG 921
QY 952 ACTGAGTTTATGCTATCTCAGAGAGAGCTGAGAAGAACTTGGTTGGAGGCAACTAT 1011
Db 922 GAGATCGTTTACGCTCAAACTGCCAAAGCTGAGAAGAACTGAAATGGAAGGCAAAATAC 981
QY 1012 GGTGTGAGGAGATGTGAGGAGCAACAAATGGAATTTGGCAAGAACAAATCCCTGGGTTAC 1071
Db 982 GGGGTAGAGGAGATGTGAGGAGCACTGTGGAATTTGGCGAGCAGACACCCCTACGGGTAT 1041
QY 1072 G 1072
Db 1042 G 1042

RESULT 9
AAA59442
ID AAA59442 standard; DNA; 1498 BP.
XX
AC AAA59442;
XX
DT 14-NOV-2000 (first entry)

```


PR	23-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	99US-0131444P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-01451192P.
PR	07-MAY-1999;	99US-0132863P.	PR	22-JUL-1999;	99US-0145115P.
PR	11-MAY-1999;	99US-0132865P.	PR	23-JUL-1999;	99US-0145216P.
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	23-JUL-1999;	99US-0145226P.
PR	14-MAY-1999;	99US-0134221P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134370P.	PR	26-JUL-1999;	99US-0145913P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145918P.
PR	19-MAY-1999;	99US-0134941P.	PR	27-JUL-1999;	99US-0145919P.
PR	20-MAY-1999;	99US-0135124P.	PR	27-JUL-1999;	99US-0145951P.
PR	21-MAY-1999;	99US-0135353P.	PR	28-JUL-1999;	99US-0146386P.
PR	24-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146388P.
PR	25-MAY-1999;	99US-0136021P.	PR	02-AUG-1999;	99US-0146389P.
PR	27-MAY-1999;	99US-0136392P.	PR	03-AUG-1999;	99US-0147038P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147204P.
PR	01-JUN-1999;	99US-0137222P.	PR	04-AUG-1999;	99US-0147302P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	09-AUG-1999;	99US-0147935P.
PR	14-JUN-1999;	99US-0139119P.	PR	10-AUG-1999;	99US-0148171P.
PR	16-JUN-1999;	99US-0139452P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139453P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-01493368P.
PR	18-JUN-1999;	99US-0139455P.	PR	16-AUG-1999;	99US-01493368P.
PR	18-JUN-1999;	99US-0139456P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139457P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139460P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139462P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139463P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139750P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139763P.	PR	26-AUG-1999;	99US-0150884P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151065P.
PR	21-JUN-1999;	99US-0139899P.	PR	27-AUG-1999;	99US-0151066P.
PR	22-JUN-1999;	99US-0139899P.	PR	27-AUG-1999;	99US-0151080P.
PR	23-JUN-1999;	99US-0140353P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140354P.	PR	31-AUG-1999;	99US-0151438P.
PR	24-JUN-1999;	99US-0140695P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-014			

PR 99US-0134941P.
PR 99US-0134700E.
PR 18-MAY-1999;
PR 19-MAY-1999;

PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139464P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142377P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144322P.
PR 20-JUL-1999; 99US-0144864P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149446P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 23-AUG-1999;	99US-0149923P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 25-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 30-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 28-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 08-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157533P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157855P.
PR 18-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 18-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 18-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146390P.	PR 28-OCT-1999;	99US-0161922P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		

Query Match 32.7%; Score 428.4; DB 3; Length 1286;
Best Local Similarity 65.1%; Pred. No. 6.3e-100;
Matches 665; Conservative 0; Mismatches 351; Indels 6; Gaps 2;

QY	51	ACACATTCGTCACCGGTGTCGGTTTCATTGGCACCCACACCGTCGTTCACTTCT	110
DB	90	AAACGTTCTAGTAACGCGGTGTCGGATACATCGGTAGTCACACGGGTCTCACTTCT	149
QY	111	CAAGCTGGCTCAGCGTTTCAATAATCGCAATTTCCATAACCTCCGTCATGGAAGCAGT	170

150 TCTCGGTGTTACTCGTGTAGTTGTTGATAAATCTCGATAATCTCCGCGGTCTCT 209
 171 GGACCGCTCGGCAAGTGGTTGGCCCTCTGTTCTCAGAACCCTCCAAATTCACCCAGG 230
 210 TCAACGTGTAAGAAA---CTGCGCGCTGAACCGGTGAACCGCTCTCTCTCCACAGGT 266
 231 CGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAACAAACATTTGATGCCGT 290
 267 GGATCTTCGAGACAGTCTGCTCTTGAGAAGATATTTCTCAGAAAACAAAGTTGATGAGT 326
 291 GATCCACTTGTGCTTGAAGCGTCTGTAAGCGTTCGAAGCGTTCGAAGCGCTGCTATT 350
 327 GATACATTTGTGGAACCTTAAGACAGTTCGTAAGAGTGTGAGAGAGCTTTGCTCTATA 386
 351 TGATTTTAATTTGGTTGGCACCACCTCAACCTCTACGAGTTTATGGCAAAAGTATAATGCAA 410
 387 TAATAATATCTGTTGTAATATCTCTCTTTTGGAGTTATGGCTCAACATGGCTGCAA 446
 411 AAGATGGTTTCTCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGA 470
 447 AATCTTGTGTTTTCATCATCAGCACTGTGTATGCTCTCCAAAGAAAGTTCTTGTATAC 506
 471 GGAGGATTTCAAGTTCAAGCTATGATCCCTATGAGCGGCAACAGCTTTTCTCGGAAGA 530
 507 AGAGGATTTCCAAATTTCTGTTTGAACCATATGAGCGAACAGATTTATCATTTAGGA 566
 531 AATGCGGAGATTTCAAGAACTGAAACAGAAATGGAAGATCATATTTCTGAGATATT 590
 567 AATCTCGCGATGATATACGTTCTGACCCGGAATGGAAGATTTATTTGCTTTAGTATT 626
 591 CAATCAGTTGGGCTCATGAAGTGCACAACTCGGTGAAGATCCAAAGGATCCCAAA 650
 627 CAATCTGTTGGGCAATCTTAGTGTGACATTTGGTGAAGACCTCGGTGTTTCCAAA 686
 651 TAACCTCATGCTTACATTCAGCAAGTAGCTGTGGAAGATTTGACTGAACCTCAATGTATA 710
 687 CAATCTCATGCTTTTGTATCAACAGTCTGTTGTTAGGAGACCTCATCTCAGTGTCT 746
 711 CGGTATGATTTCCAAAGGATGCTCTGATCCGGGACTATATCCATGTATGATGA 770
 747 TGAAGATGATTAACAATCAAGATGAAACAGGGTTGAGATTTACATTTGATTTGA 806
 771 CTGGCAGATGGCCATTTGCTGCCCTGCGAAAGCTTTTCACACGGAAGAACATAGTTG 830
 807 TTTAGCTGATGGACACATAGCGCTCTACGTAAGC---TAGAAGATTTGCAAGATCGGTTG 863
 831 TACTGCTTCAACCTGGAACTGGTCTGTAACATCTGCTTGAATGTTTACAGCAT 890
 864 TGAAGTGATCAATCTCGGAACAGGAATGGAACATCAGTTCTTGAATGGTTGATGCTTT 923
 891 TGAAGAGGCTTCTGGCAAGAAAATTCAGTAAATTTATGTCCAAAGAACCGGAGATGC 950
 924 TGAAGAGCATCCGGAAGAAATCCCTCTGTTGATCGCGGACGTCGACCGGAGATGC 983
 951 GACTGAGTTTATGCACTACAGAGAGCTGGAAGAACTTGGTTGGAAGGCAACTA 1010
 984 TGAAGTTGTATACGCTCGACGGAAGAGCAGAAAGTGAATTTGAATTTGAAGGCCAAATA 1043
 1011 TGTGTGAGGAGATGTGACGGGACCAATGGAATTTGGCAAGAAACAAATCCCTGGGGTTA 1070
 1044 TGGATTTGAGAGATGTGAGATTTTGTGGAATCTGGGCTAGTAATATCCCTTATGATA 1103
 1071 CG 1072
 1104 TG 1105

RESULT 14
 ID AAA59435
 XX AAA59435 standard; DNA; 479 BP.
 AC
 XX

14-NOV-2000 (first entry)
 XX Nucleotide sequence of UDP-galactose 4-epimerase.
 XX
 XX UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
 XX raffinose pathway; raffinose; stachyose; carbohydrate metabolism;
 XX starch level; ss.
 XX
 XX Glycine max.
 XX
 XX W0200047755-A2.
 XX
 XX 17-AUG-2000.
 XX
 XX 09-FEB-2000; 2000WO-US003453.
 XX
 XX 10-FEB-1999; 99US-0119588P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Carlson TJ, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;
 XX Rafalski JA, Thorpe CJ;
 XX WPI; 2000-549152/50.
 XX P-PSDB; AAB07815.
 XX
 XX Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant
 XX cells to e.g. decrease cell wall constituents and increase starch
 XX content.
 XX
 XX Claim 2; Page 44; 61pp; English.
 XX
 XX The present sequence encodes a UDP-galactose 4-epimerase. The enzyme is
 XX responsible for the interconversion of UDP-glucose and UDP-galactose. As
 XX the activity of the enzyme appears to be particularly limiting to carbon
 XX flux into the raffinose pathway, reduction of its activity should
 XX decrease the levels of raffinose and stachyose in seeds. The
 XX polynucleotide is useful to transform plant cells to enhance the degree
 XX of expression of UDP-galactose-4-epimerase within these cells. Changes in
 XX the expression of UDP-galactose-4-epimerase within a cell enables
 XX modification of plant carbohydrate metabolism. This allows growth of
 XX grains with reduced cell wall constituents (fiber) and increased levels
 XX of starch
 XX
 XX Sequence 479 BP; 120 A; 115 C; 104 G; 125 T; 0 U; 15 Other;
 XX
 XX Query Match 32.4%; Score 424.8; DB 3; Length 479;
 XX Best Local Similarity 94.1%; Pred. No. 3.6e-99;
 XX Matches 443; Conservative 0; Mismatches 27; Indels 1; Gaps 1;
 QY 9 CACTTCTCTCCCTCTTATTGCAAGCATGGTGTCTTCTCCCAACACATTTGTCACCG 68
 Db 3 CAATTTCTCTCCCTCNCNAATGCAAGTGGGNGTCTCTCCCAACANATTTGTCACCG 62
 QY 59 TGGTCCCGGTTTCAATGGCACCCACACCGTCTCAGTTCTCAAAGCTGGCTTCAGCGT 128
 Db 63 NGNGGGCGGTTTCAATGGNACCACACCGTCTGTTCACTTCTCAAAGCTGGCTTCAGCGT 122
 QY 129 TTCAATATCGCAATTTGATTAATCCGTCATGGAAGCATGGAACCGCTCCGCCAAGT 186
 Db 123 TTCAATATCGCAATTTGATTAATCCGTCATGGAAGCATGGAACCGCTCCGCCAAGT 182
 QY 189 GGTGGGCGCTCTGTTTCTCAGAACCTCAATTCACCCAGCGCATCTCCGGAATAGGA 248
 Db 183 GGTGGGCGCTCTGTTTCTCAGAACCTCAATTCACCCAGCGCATCTCCGGAATAGGA 242
 QY 249 TGACTTGGAGAACTCTTCTCCAAACAAACATTTGATGCCGTGATCCACTTTGCTGCTT 309
 Db 243 TGACTTGGAGAACTCTTCTCCAAACAAACATTTGATGCCGTGATCCACTTTGCTGCTT 302
 QY 309 G-AAAGCGGTTGCTGAAGCGTTTCGAAGCGCGCTGCTTATTTGATTTTAAATTTGCTG 367
 Db 303 GAAAGCGGTTGCTGAAGCGTTTCGAAGCGCGCTGCTTATTTGATTTTAAATTTGCTG 362

PR	10-AUG-1999;	99US-0148171P.	Db	90	AAACGTTCTAGTAAGCGTGGTCTGGATACATCGGTAGTCACACGGTGTCTCAACTTCT	149
PR	11-AUG-1999;	99US-0148319P.	Qy	111	CAAAAGCTGGCTTCAAGCGTTTCAATAATCGACAAATTCGATTAATCTCGTATGGAACAGT	170
PR	12-AUG-1999;	99US-0148341P.	Db	150	TCTCGGTGGTTACTCCGTCGTAGTGTGTTGATAATCTCGATAATTTCTCCGCGTCTCTCT	209
PR	13-AUG-1999;	99US-0148565P.	Qy	171	GGACCGGTCGCGCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCAATTCACCCAGGG	230
PR	16-AUG-1999;	99US-0149175P.	Db	210	TCACCGTGTGAAGAACTCGCG--CTGAACACGGAGAACGCTCTCTCTCCACCAAGT	266
PR	18-AUG-1999;	99US-0149426P.	Qy	231	CGATCTCCGGAATPAGGATGACATTTGAGAGAACTCTTCTCAAAACAAACATTTGATCCGT	290
PR	20-AUG-1999;	99US-0149722P.	Db	267	GGATCTTCGAGACAGGTCTGCTCTTGAGAGAGATATTCAGAACAAAGTTGATCAGT	326
PR	23-AUG-1999;	99US-0149902P.	Qy	291	GATCCACTTTGCTGGCTTGAAGCGGTGCTGAAGCGTTGCGAACCGTTCGAAACCGCTCGCTATT	350
PR	25-AUG-1999;	99US-0150566P.	Db	327	GATACATTTGCTGGACTTAAGCAGTTGGTAGAGTGTAGAGAGCCCTTGGCTTTATTA	386
PR	26-AUG-1999;	99US-0150884P.	Qy	351	TGATTTTAATTTGGTGGCACCATCAACCTCTACGAGTTTATGGGAAAGTATTAATTCGAA	410
PR	27-AUG-1999;	99US-0151065P.	Db	387	TAATAATAATCTTGTGGTACTATTACTCTTTTGGAAAGTTATGGCTCAACATGGCTGCAA	446
PR	01-SEP-1999;	99US-0151930P.	Qy	411	AAAGATGGTTTCTCATCATCTGCAACCGTTTATGGCCCAACCTGAAAGATACCGTGTGA	470
PR	07-SEP-1999;	99US-0152363P.	Db	447	AAATCTGTGTTTCAATCATCAGCACTGTATAGCTCTCCAAAGAGTTCCTTGTGTAC	506
PR	13-SEP-1999;	99US-0153758P.	Qy	471	GGAGGATTTCAAGTTACAAGTATGAATCCCTATGAACGACCAAGCTTTTCTCGAAGA	530
PR	15-SEP-1999;	99US-0154018P.	Db	507	AGAGGAGTTTCCAATTTCTGTTGAACCATATGGACGCAACAAAGCTATTCATTGAGGA	566
PR	20-SEP-1999;	99US-0154458P.	Qy	531	AATTGCCGAGATATTCAGAACTCAACAGATCGAAGATCATATTTACTGAGATACTT	590
PR	22-SEP-1999;	99US-0154779P.	Db	567	AATCTGCGCGATGTATACGTTCTGACCCGGAATGGAAGATATATTTGTTAGGTATTT	626
PR	04-OCT-1999;	99US-0157117P.	Qy	591	CAATCCAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGCAATCCCAA	650
PR	06-OCT-1999;	99US-0157865P.	Db	627	CAATCTGTGGTGACATCTAGTGGTGACATTTGGTGAAGACCTCTGTTGTTATTCAAA	686
PR	07-OCT-1999;	99US-0157865P.	Qy	651	TAACCTCATGCTTACATTCAGCAAGTGTGTGGAAGATTGACTGAACTCAATGTATA	710
PR	08-OCT-1999;	99US-0158232P.	Db	687	CAATCTCATGCTTGTGTACAAAGTGGTGTGTGTAGAGACCTCATCTCACTGTCTT	746
PR	12-OCT-1999;	99US-0158269P.	Qy	711	CGGTGATGATTATCCACGAGGATGGCTCTGCGATCCGGGACTATATCATCTGTATGGA	770
PR	13-OCT-1999;	99US-0158293P.	Db	747	TGGAATGATTACATACAAAGATGGAACAGGGGTTTCGAGATTACATTCATGTTATGA	806
PR	14-OCT-1999;	99US-0159331P.	Qy	771	CTTGGCAGATGGCCATATTTGCTGCCCTGGAAGCTCTTCAACGAGAGACATAGGTTG	830
PR	14-OCT-1999;	99US-0159637P.	Db	807	TTTAGCTGATGGACACATAGCGGCTCTACGTAAAGC--TAGAAGATTGCAAGATCGGTTG	863
PR	18-OCT-1999;	99US-0159838P.	Qy	831	TACTGCTTACAACCTGGAACTGGTCTGGAACATCTGTGCTTGAATGTTTACAGCATT	890
PR	21-OCT-1999;	99US-0160741P.	Db	864	TGAAGTGTACATCTCGAACAGGAATGGAACATCAGTTCTTTGAAATGTTGATGCTTT	923
PR	21-OCT-1999;	99US-0160770P.	Qy	891	TGAAAAGCTTTGGCAAGAAATTCAGTAAAAATTTATGTCGAAGAGACCCGGGAGATGC	950
PR	22-OCT-1999;	99US-0160981P.	Db	924	TGAAAAGCATCCGGAAGAAAATTCCTCTGTGTGATCGCGGACGTCGACCGGAGATGC	983
PR	25-OCT-1999;	99US-0161404P.	Qy	951	GACTGAGTTTATGTCATCTACAGAGAGCTGGAAGAACTTGGTTGGAAGGCAAACTA	1010
PR	25-OCT-1999;	99US-0161405P.	Db	984	TGAAGTTGTATACCTCGACGGAAGAGAGAGAAAGTGAATTCGAATTTGGAAGGCCAAATA	1043
PR	26-OCT-1999;	99US-0161359P.	Qy	1011	TGGTGTGGAGGAGATGTGACGGACCAATGGAAATTTGGCAAGAGAACAAATCCCTGGGGTTA	1070
PR	26-OCT-1999;	99US-0161361P.	Db	1044	TGGGATTTAGGAGATGTGTAGAGATTTGTGGAACCTGGGTAGTATATATCTTTATGATA	1103
PR	28-OCT-1999;	99US-0161920P.	Qy	1071	CG 1072	
PR	28-OCT-1999;	99US-0161993P.	Db	1104	TG 1105	

Search completed: March 22, 2004, 17:46:07
Job time : 578 secs

Query Match 32.3%; Score 423.6; DB 3; Length 1292;
Best Local Similarity 64.8%; Pred. No. 1.1e-98;
Matches 662; Conservative 0; Mismatches 354; Indels 6; Gaps 2;
QY 51 ACACATCTGGTCACCGGTGGTGGCGGTTTCAATGGCACCCACACCGGTGTTCACTTCT 110

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 17:35:58 ; Search time 113 Seconds
(without alignments)
6443.324 Million cell updates/sec

Title: US-09-913-064A-13
Perfect score: 1312
Sequence: 1 gcacagcgcactctctctccc.....aaaaaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCITUS COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261.8	20.0	100848	4	US-09-596-002-39
C 2	261.6	19.9	13206	4	US-08-961-527-33
3	258.6	19.7	1131	4	US-09-540-236-260
C 4	247.2	18.8	3140	4	US-08-961-527-158
5	244	18.6	1488	4	US-09-023-655-880
C 6	241.4	18.4	1830121	4	US-09-557-884-1
7	241.4	18.4	1830121	4	US-09-643-980A-1
C 8	240.8	18.4	1249	3	US-09-113-536-1
9	240.8	18.4	1249	4	US-09-624-183-1
10	240.8	18.4	1249	5	PCR-US95-05785-1
11	239	17.5	1038	4	US-09-543-681A-776
12	199	15.2	1092	4	US-09-489-039A-6024
13	158.2	12.1	307	4	US-09-313-294A-7524
C 14	146	11.1	4597	4	US-09-221-017B-1093
15	144.8	11.0	381	4	US-09-424-311-1
16	118	9.0	1026	4	US-09-134-000C-2119
17	114.8	8.8	990	4	US-09-134-000C-2833
18	106.4	8.1	1410	4	US-09-107-532A-2282
19	97.6	7.4	351	4	US-09-424-311-2
20	85	6.5	1026	4	US-09-328-352-3204
21	64.2	4.9	1035	4	US-09-489-039A-916
C 22	62.4	4.8	907	4	US-08-961-527-311
23	53.8	4.1	580073	4	US-08-545-528D-1
24	50.8	3.9	222	3	PCR-US93-00869-15
25	50.8	3.9	222	5	PCR-US93-00869-15
C 26	48.4	3.7	7218	1	US-08-232-463-14
27	48.4	3.7	2649	2	US-08-718-964-1

Sequence 1, Appli
Sequence 1, Appli
Sequence 17, Appli
Sequence 17, Appli
Sequence 17, Appli
Sequence 14, Appli
Sequence 14, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 359, App
Sequence 7, Appli
Sequence 24, Appli
Sequence 313, App
Sequence 3, Appli

28 48.4 3.7 2649 2 US-09-059-964A-1
29 48.4 3.7 2649 2 US-08-842-341-1
30 46.8 3.6 2619 4 US-08-983-502-17
31 46.8 3.6 2619 4 US-09-516-747-17
32 46.8 3.6 2619 5 PCT-US96-10521-17
33 46.8 3.6 2887 4 US-08-983-502-14
34 46.8 3.6 2887 4 US-09-516-747-14
35 46.8 3.6 2887 5 PCT-US96-10521-14
36 46.4 3.5 2322 1 US-08-618-164-1
37 46 3.5 1664976 4 US-08-916-421B-1
38 44.8 3.4 4526 1 US-07-855-412B-4
39 44.8 3.4 4526 2 US-08-308-887A-4
40 44.8 3.4 4526 3 US-08-881-094-4
41 44.4 3.4 1044 4 US-09-543-681A-359
42 43.6 3.3 1558 1 US-08-455-550-7
43 41.6 3.2 1866 4 US-09-673-395A-24
44 41.4 3.2 577 4 US-09-149-476-313
45 41 3.1 6124 4 US-08-213-419B-3

US-09-596-002-39
; Sequence 39, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE OF INVENTION: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 39
; LENGTH: 100848
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 39
; PUBLICATION INFORMATION:
US-09-596-002-39

Query Match 20.0%; Score 261.8; DB 4; Length 100848;
Best Local Similarity 54.7%; Pred. No. 4.7e-64;
Matches 571; Conservative 0; Mismatches 457; Indels 15; Gaps 2;

QY 31 AGCATGGTCTCTCCCAACACATCTGGTCACCGGTGGTCCGGTTTCATTGGCACC 90
Db 51531 AGCGAGCATTTATTATCCAAAAAATCTGGTAACAGCGGAGCGGGTTATATCGGTCA 51590
QY 91 CACACCGCTGGTTTCAGCTTCTCAAGCTGGCTTCAGCGTTTCAATAATCGACAATTCGAT 150
Db 51591 CATACTTTGATTGAGCTGTACAGCGGGATTGTGCGACGTGGTGTATGATAATTTATCC 51650
QY 151 AACTCCGTCTAGGAAGCAGTGGACCGGTCCGCCAAGTGGTGGCCCTCTGCTTTCTCAG 210
Db 51651 AATTCAAGTTTGTGTGAGTGCAGCGAGTTGAGCAAAATTTGTCACCAACATATG----- 51705
QY 211 AACCTCAATTCACCCAGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCC 270
Db 51706 -----AATTATCAAGTGAATTTTGGATAAACGCATCTTGACGCTGTTTTAAG 51758
QY 271 AAAACAACATTTGATGCCGTGATCCACTTTGCTGGTTGAAAGCGGTTCGCTGAAGCGTT 330
Db 51759 GCATCATCAGTTTTTTGCTGTAGTTCATTTTGCAGGGCTAAAGCTGTGGGAGAAAGCACC 51818

ALIGNMENTS

RESULT 1
US-09-596-002-39
; Sequence 39, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE OF INVENTION: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 39
; LENGTH: 100848
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 39
; PUBLICATION INFORMATION:
US-09-596-002-39

```
QY 331 GCGAAGCCCGTCGCTATTTGATTTTAAATTTGGTTGGCACCATCAACCTCTACGAGTTT 390
Db 51819 AAAAATCCCTTAAATATATCAAAATTAATGTTGCGGCACACTCAGTCTGTAGAGCTG 51878
QY 391 ATGCGAAGATATAATGCGAAGATGTTTCTCATCTGCAACCGTTATGGCCAA 450
Db 51879 ATGCGAAGATATGGGGTTAAATTTGCGTTTTCATCATCGCAACCGTTATGGCAGC 51938
QY 451 CTGAAAAGATACCGTGTGAGGAGGATTTCAAGTTTCAAGCTATGAATCCCTATGGACGG 510
Db 51939 TCAATTCGCTACCAATCACCAGGATATGCGCGTTTTTGTACAGTCTTATGGTCAA 51998
QY 511 ACCAAGCTTTTCTGGAGAAATTCGCCGAGATATTCGAAAGCTGCAACCAAGATGGAAG 570
Db 51999 AGTAATGATGGTACAGCATATATGGAGGATTTGGTGAATGCTGATGATACATGGAAAT 52058
QY 571 ATCATATTACTGAGTACTTCAATCCAGTTGGGGCTCATGAAGTGGCAAACTCGGTGAA 630
Db 52059 GTTGCTGTTGAGGTATTTTAAATCCAAATGGTGACATGAAGTGGCGGAATGGTGAG 52118
QY 631 GATCCCAAGGGATCCCAATTAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAGA 690
Db 52119 GATCCCAAGCGATATTCCTCAATTAACCTGATGCCATACATCAGTCAAGTGGCGGTAGGTAAT 52178
QY 691 TTGACTGAATCAATGATACGCTCATGATTAATCCAAAGGAGGATGCTCTGCGATCCGG 750
Db 52179 CTTAAACAGCTGCTGTATTTGGCAATGATATGAGACACCTGACGCGTACGGGGTGGCA 52238
QY 751 GACTATATCCATGTGATGACTTTGGCAGATGGCCATATTTGTCGCTCGGAAA---GCTC 807
Db 52239 GATTATATTATCATGTGTGATTTGGCAAAAGGTCATGTGGCAGCACTGCATTTATTTGGTT 52298
QY 808 TTCACAAAGGAGAAATGTTGATGCTTCAACCTGGGAACCTGGTGGGAACATCT 867
Db 52299 GGGCAATCTGTGGCATTGATTTGTCCTCATCACTTGGCAGAGTCAAGGAACCTTCG 52358
QY 868 GTGCTTGAATGTTACAGCATTTGAAAGGCTTTCTGCAAGAAATTTCCAGTAAATTA 927
Db 52359 GTTTTACAGCTTATTAAGCGTTTGAGTCAAAATCTGCTCAAGCGTACCGTATGTTATT 52418
QY 928 TGCTCAAGAGACCGGAGATGCGTACGAGTTTATGATCTACAGAGAGAGCTGAGAAA 987
Db 52419 ACCAGCAGACGACAGTGACATTTGAGCGCTTATGCGCTGATTAAGCCAAAGCT 52478
QY 988 GAATCTGTTGGAAGGCAAACTATGCTGTGAGGAGATGTGAGGAGCAAAATGGAATGG 1047
Db 52479 ATACTTAATCGACACAGAGCTTGATATTGAGCGTATGTGCGTAGATACATGGCGGTG 52538
QY 1048 GCAAGAACAAATCCCTGGGGTTA 1070
Db 52539 CAGTCAAAATATCCACACGGCTA 52561
```

RESULT 2

```
US-08-961-527-33/c
; Sequence 33, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13206 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-33
```

```
Query Match 19.9%; Score 261.6; DB 4; Length 13206;
Best Local Similarity 55.3%; Pred. No. 1.9e-64;
Matches 560; Conservative 0; Mismatches 434; Indels 18; Gaps 2;

QY 50 AACACATCTGTCACCGGTGGTCCGGTTTCATTTGSCACCCACACCGCTTCAGCTTC 109
Db 9213 AAAAGATTTTGGTAACTGTTGGTGGTCCGGTTTATTTGGAACCCACACTGTTATTTGAATTGA 9154
QY 110 TCAAGCTGTCCTCAGCGTTTCAATTAATCGACAATTCGATTAACCTCCGTCATGGAAGCAG 169
Db 9153 TCCAAGCAGGTATCAAGTTGTTGGTGGTAATCTTCTCAATAGCAATCGTAAGAGTT 9094
QY 170 TGAACCGCTCCGCCAAGTGGTGGCCCTCTGCTTCTCAGAACCCTCCAAATTCACCCAGG 229
Db 9093 TAGAAGTTGTTGAAGAAATCAGAGGAGTTGAGATTCCT-----TTCTATGAGG 9046
QY 230 GCATCTCCGGAATAGGATGACTTGGAGAAACTCTTCTCCAAACAAACATTTGATGCGG 289
Db 9045 CAGATATTCGTGACACAGATACCTCAGAGATATTTTCAAGCAGGAGAAACCGACTGGTG 8986
QY 290 TGATCCACTTTGCTGGCTTGAAGCCGTTGCTGAAAGCGTTGCGAGAGCCCGTCGCTATT 349
Db 8985 TCATTCACCTTTGCTGTTGAGGCTGTTGGCGAATCAACAGTATCCCTCTTGGCTACT 8926
QY 350 TTGATTTTAATTTGGTTGGCACCATCAACCTCTACGAGTTTATGCGCAAAAGTATAATTGCA 409
Db 8925 ATGCAACAATATCGCTGGAACTGTGAGCTTTTGAAGCCATGGAAGAAACAACTGTA 8866
QY 410 AAAAGATGTTTCTCATCATCTGCAACCGTTTATGCGCAACCTGGAAGATACCGTGTG 469
Db 8865 AAAACATCATCTTCACTTCTCTGCGACAGTTTACGGGGATCCGACACAGTCCCATCT 8806
QY 470 AGGAGGATTTCAAGTTTACAGCTATGAATCCCTTATGAGGAGCAACAGCTTTTCTCGAAG 529
Db 8805 TGGAGATTTTCCACACTTTCAGTGACCAACCCATACGCTCGTACTAAGCTCATGTAGAGG 8746
QY 530 AAATGCGCGAGATATTCAGAAAGCTGAACAGATGGAAGATCATATTTACTGAGTACT 589
Db 8745 AAATTTGACTGATATTTTCAAAAGCAGACTCAGATGGAATGTTGTCTTCTCGTTACT 8686
QY 590 TCAATCCAGTTGGGCTCATGAAGTGGCAAACTCCGTTGAAGATCCCAAGGGCATCCCA 649
Db 8685 TTAACCAATCGAGAGCCATGAGTGGTGTGATTTGGAGAGAAATCCAAACGGTATTTCCAA 8626
QY 650 ATAACTCATGCTTACATTCAGCAAGTAGCTGTTGGAAGATTTGACTGAACCTCAATGTAT 709
Db 8625 ACAATCTTGGCCATATGTGACTCAAGTAGCCGTTGGAAAAATTTAGAGCAAGTCAAGTGT 8566
QY 710 ACGTCAATGATTTCCAAAGGAGGATGGCTCTCCGATCCGGGACTATATCCATGTGATGG 769
Db 8565 TTGAGAGCATTAAGATACGGAAGATGGAACAGGTGTTGCTGACTATATCCACGTTGTGCG 8506
```


REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 158:
SEQUENCE CHARACTERISTICS:
LENGTH: 3140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-158

Query Match 18.8%; Score 247.2; DB 4; Length 3140;

Best Local Similarity 55.8%; Pred. No. 1.2e-60;
Matches 539; Conservative 0; Mismatches 423; Indels 18; Gaps 2;

QY 55 ATTCTGGTCCAGCGTGGTGGCGGTTTCATTGGCACCACCCAGCGTGGTTCAGCTTCTCAAA 114
Db 965 ATATTGGTAACAGCGGAGCTGGTTATATTGGTAGCCATACCGTAGTAATAAT 906
QY 115 GCTGGCTTCAGCGTTCAATATCGACAAATTCGATACTCGTATGGAAGCAGTGGAC 174
Db 905 TTAGGAAGAGGATATATATTCGACAACTTTCGAACTTAGCATCTTGGTGTAGAC 846
QY 175 CGGTCGCGCAAGTGGTGGCGCTGCTTCTCGAACCTCCAAATTCACCCAGGCGAT 234
Db 845 CGTATTGAACGATATACAGATACGTCC-----CGTGTTCAGAAATAGAT 798
QY 235 CTCGGGAATAGGATGACTTGGAGAACTCTCTCCAAACAAATTTGATCCGTGATC 294
Db 797 GTTGTGATAAACAAGCATTTGAAAGATTTTGAACAGAAATCGATTGATGCTCAAT 738
QY 295 CACTTGTGCTGTTGAAAGCGTGTGAAAGCGTTGCGAAGCGTGGCGCTGCTATTTGAT 354
Db 737 CATTTGCGAGTTATTAAGCTGTGCGAGAAATCGTGCAAGCCGTGATGTATCAAAA 678
QY 355 TTPTAATTTGGTGGCCACCATCACTCTACGAGTTTATGGCAAGTAAATTTGCAAAAAG 414
Db 677 AATAATATTATGATGATCACTTGGCACTTGTGGAAGTATGTCAGAATTTAATGTTAAAAG 618
QY 415 ATGGTTTCTCATCTCTGCAACCGTTTATGCGCAACCTGGAAGATACCGTGTGAGAG 474
Db 617 ATTGCTTTTCTCAAGTCGAGTGTATATGGAATTAACAATCAGTCACCTCTAATTTAG 558
QY 475 GATTTCAAGTTACAAGTATGATCCCTATGGACGACCAAGCTTTTCTCGAAGAAAT 534
Db 557 ACGATCAACAAGTCGCAAAATCCTTATGGGTATACGAAGTATGCTTGAGCAAT 498
QY 535 GCCCGAGATATTCAGAAAGCTGAAACGAAATGGAAGATCATATTACTGAGATCTCAAT 594
Db 497 TTAAAGATGTTTCATGTGGCAGATTTCAGAAATGAGTATTCGCTTCTGTTTCAAT 438
QY 595 CCAGTTGGGCGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGATCCCAATTAAC 654
Db 437 CCAATTGGTGCATGAGTCTGGTTGATGGGGAAGATCCCTCAGAAATCTTAACAC 378
QY 655 CTCATGCCCTTACATTGAGCAAGTGTGTTGGAAGATTCAGTGAATCAATGATATACGGT 714
Db 377 TTGATGCCCTTATATTCACAAGTAGCGGTAGTAAAGCTATCAGAGCTAAGTGTGTTCGA 318
QY 715 CATGATTATCCAAAGGAGGATGGCTCTGGATCCGGACTATATCCATGTGATGAGCTTG 774
Db 317 AATGATTATGATACGCTTATGATGGTACTGTGTGGTGTGATATATCCATGTAGTATTA 258
QY 775 GAGATGGCCATATTGCTGCCCTGGGAAAGCTTTTCAACACGGAACATAGTTGACT 834
Db 257 GCAATAGGACATATAAAGCTTTAGAAAAGT-----ATCTGAAAAACAGATGTTTAT 204
QY 835 GCTTACAACTCGGAACTGGTGTGGAACATCTGTGCTTGAATGTTTACAGCATTTGAA 894
Db 203 ATCTATTAACCTAGCTCGGAGAGGCACTAGTGTATTATACAACTTGTAAATACATTGAA 144

RESULT 5

US-09-023-655-880
Sequence 880, Application US/09023655
Patent No. 6607879

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.

APPLICANT: Susan G. Stuart

APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

TITLE OF INVENTION: EXPRESSION

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESS: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,655

FILING DATE: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 880:

SEQUENCE CHARACTERISTICS:

LENGTH: 1488 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENEANK

CLONE: g1119216

US-09-023-655-880

Query Match 18.6%; Score 244; DB 4; Length 1488;

Best Local Similarity 54.3%; Pred. No. 6.6e-60;
Matches 565; Conservative 0; Mismatches 460; Indels 15; Gaps 3;

QY 56 TTCTGCTACCGGTGGTGGCGGTTTCATTGGCAACCCACACCGGTCGTTTCTCAAG 115
Db 107 TGCTGGTAAACAGGTGGGCTGGCTACATTGGCAGCCACACCGTGTGGAGCTGCTGAGG 166
QY 116 CTGGCTTCAGCGTTTCAATTAATCGAATTCGTAATCCGTCATGGAGCAGTGGACC 175
Db 167 CTGGCTACTTGGCTGTGGTTCATCGATACTTCCATTAATGCTTCCGTGAGGGGGCTCCC 226

176 GCGT-----CCGCCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCAAATTCACCCAGG 229
Db TGCCTGAGAGCTTCGCGCGCTCCAGAGCTGACAGCCCGCTCTGTGGATTTGAGGAGA 286
Qy GCGATCTCCGGAATAGGAGTACTTGGAGAACTCTCTCCAAAACAAATTTGATGCCG 289
Db TGGACATTTGGACAGGAGCCCTACAGCGTCTCTCAAAAAGTACAGCTTTATGCCG 346
Qy TGATCCACTTTGCTGGCTTGAAGCGTTGCTGAAGCGTTGCGAAGCCCGCTGCTATT 349
Db TCATCCACTTTGCGGGCTCAAGCCCGTGGCGAGTCGGTGCAGAAAGCTCTCGATTAT 406
Qy TTGATTTTAAATTTGTTGGACCATCAACTCTACGAGTTTATGGCAAGTATAATTGA 409
Db ACAGATTTAACTGACCGGACCATCAAGCTTCTGGAGATCATAAGGCCACCGGGTGA 466
Qy AAAAGATGGTTTCTCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTG 469
Db AGAACCTGGTGTTCAGCAGCTCAGCCACTGTGACGGAAACCCCACTACCTGCCCTTG 526
Qy AGGAGG--ATTTCAGCTTACAGCTATGATCCCTATGACGAGCCACAGCTTTCTCTG 526
Db ATGAGGCCACCCACCGGTGGTTGTACCAACCTTACGCAAGTCCAAAGTTCTTATCG 586
Qy AAGAAATTTGCCCGAGATATTTCAGAAAGCTGAACAGAAATGGAAGATCATATTACTGAGAT 586
Db AGAAATATGATCCGGGACCTGTGCCAGCAGACAAAGCTTGGAAAGTAGTCTGCTCGCT 646
Qy ACTTCATCCAGTTGGGGCTCATGAAGTGGCAACTCGGTGAAGATCCCAAGGGATCC 646
Db ATTTCAACCCCAAGSGTCCCATGCTCTGCGTGCATTTGTGAGGATCCCAAGGGATAC 706
Qy CAATAACCTCATCCCTTACATTCAGCAAGTAGTGTGGAAGATTGACTGAACCTCAATG 706
Db CCACAACTCATGCTTATGCTCTCCAGTGGCGATCGGCGACGGAGCCCTGATG 766
Qy TATACGCTCATGATATCCAAAGAGGATGGCTCTGCGATCCGGACTATATCCATGTGA 766
Db TCTTTGGCAATGACTATGACACAGAGATGGCACAGGTGTCCGGGATATACATCATGTG 826
Qy TGGACTTGGCAGATGGCCATATGCTCCCTGCGAAAGCTCTTCAACAGGAGAACATAG 826
Db TGGATCTGGCAAGGGCCCATTCGACGCTTAAGGAGCT-----GAAGACAGTGTG 880
Qy GTTGTACTGTCTTCAACCTGGGAAGTGTGCTGGAACATCTGTCTTGAATGTTGATACAG 886
Db GCTCCCGGATCTCAACCTGGCAGCGGCACAGGCTATTCAAGTGTGCGAGATGCTCCAG 940
Qy CATTTGAAGAGGCTTCTGGCAAGAAATTCAGTAAATTTATGTCGAAGAGCCGGAG 946
Db CTATGGAGAGGCTCTGGGAAGAGATCCCGTACAAAGTGTGGCAGCGGGGAAGGTG 1000
Qy ATGCGACTGAGGTTTATGCACTACAGAGAGCTGAGAAAGAACTTGGTTGAAGGCAA 1006
Db ATGTGGCAGCTGTTACGCCAACCCAGCTGCGCCCAAGAGGAGCTGGGTGGACAGCAG 1060
Qy ACTATGTGTGGAGAGATGTGAGGACCAATGGAATTTGGCAAGAACATCCCTGGG 1066
Db CCTTAGGGCTGGACAGATGTGTGAGGATCTGTGGCGCTGGCAGAGAGCAATCTTCAG 1120
Qy GTTACGCGGGGAAGCCCTTGA 1086
Db GCITTTGGCACGAAGCCTGA 1140

RESULT 6

US-09-557-884-1/c

; Sequence 1, Application US/09557884

; Patent No. 6506581

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: The Nucleotide sequence of

the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB186P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match

Best Local Similarity 18.4%; Score 241.4; DB 4; Length 1830121;

Matches 557; Conservative 0; Mismatches 442; Indels 21; Gaps 3;

Qy	54	CATTCTGCTCACCGGTGGTCCGGTTTCATTGGCACCCACACCGGTGTTGAGTCTCTCAA	113
Db	378456	CATTTTAGTGGGTGGTGGCTGTTATATCGTTCTCACACAGTTGTAGATTATTAA	378397
Qy	114	AGTGGCTTCAGCGTTTCAATATCGAATTCGATAATCCCTCATCGTATGGAAGCAGTGA	173
Db	378396	TGTTGGCAAGAGGTGGTGTATTAGATAATCTTTGCAATTCATCGCCAAAATCCCTGA	378337
Qy	174	CGCGTCCGCCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCCAATTCACCCAGGCGA	233
Db	378336	CGCGTAAACAAATTCAGGCA-----AAGACAAAGTTTATGAGGCGA	378289
Qy	234	TCTCGGAATAGGATGACTTGGAGAACTCTCTCCAAAACAAATTTGATGCGGTGAT	293
Db	378288	TATTTTAGATCGTCTTTGTACAAAATTTTTCAGAAAATAGATTAATCGGTTAT	378229
Qy	294	CACTTTCCTGGCTTGAAGCGTTGCTGAAAAGCTTGCAGAGCCCGTCCCTATTTTGA	353
Db	378228	TCACTTTCAGGGTTAAAGCCGTAGGGGAAAGTGTCAAAGGCGCAGAAATATTACAT	378169
Qy	354	TTTAAATTTGGTGGCACCCTCAACCTCTACAGTTTATGCAAAAGTATATTCGAAAA	413
Db	378168	GACATGTCGCTGGCACCTTGTATTCAAGAAATGAAAAGACGAGGTGTTGGAA	378109
Qy	414	GATGGTTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGA	473
Db	378108	CTTTGTATTAGCTCATCTGCAACGGTTTACGGGAGNCACAAAATTTATYCCANTTACAG	378049
Qy	474	GGATTTCAAGTTACAAGCTATG---AATCCCTATGGAGCGGACCAAGCTTTTCTCTGGAAGA	530
Db	378048	GGATTTGTGAAGTCGGCGGTACCAACACCCCTTATGTTACATCTAAATATATGTTGAGCA	377989

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-305

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1249 BASE PAIRS

TYPE: NUCLEIC ACID

STRANDEDNESS: SINGLE

TOPOLOGY: LINEAR

MOLECULE TYPE: cDNA

PCT-US95-05785-1

Query Match 18.4%; Score 240.8; DB 5; Length 1249;
Best Local Similarity 54.1%; Pred. No. 5e-59;
Matches 563; Conservative 0; Mismatches 462; Indels 15; Gaps 3;

QY 56 TTCTGTACCGGTGTGCGGTTTCATTTGGCAACCCACACCGTCGTTGAGTTCCTCAAG 115
DB 106 TGCTGTAACAGGTGGGGTGGCTACATTTGGCAGCCACACCGTGTGGAGCTGCTGGAGG 165
QY 116 CTGGCTTCAGCGTTTCAATAATCGCAATTCGATACTCCGTTCATGGAAGCAGTGGACC 175
DB 166 CTGGCTACTTGGCTGTGGTCATGATACTTCCATTAATGCTTCCGTGGAGGGGCTCC 225
QY 176 GCGT-----CGCCAAAGTGTGGCCCTCTGCTTCTCAGAACCTCCAAATCACCCAGG 229
DB 226 TGCTGAGAGCCTGCGCGCGGTCCAGAGCTGACAGCGCGCTCTGTGGAGTTTGGAGGA 285
QY 230 GCGATCTCGGAATAGGATGACTTCGAGAACTCTTCCAAAACCAATTTGATCGCG 289
DB 286 TGGACATTTTGGACAGGAGCCCTACAGGCTCTTCAAAAAGTACACTTATGCGG 345
QY 290 TGATCCACTTTGTGCTTGAAGCGGTGTGTAAGCGTTGCGAAGCCCTCGCTATT 349
DB 346 TCATCCACTTTGGGGGCTCAAGGCCCTGGCGAGTCCGTGCGAGAGCCTCTGGATTATT 405
QY 350 TTGATTTTATTTGTTGGCAACCATCACTCTACGAGTTTATGGCAAGTATAATTGCA 409
DB 406 ACAGAGTTAACTGACCGGACCATCCAGCTCTCGAGATCATGAAGGCCCGGGGTGA 465
QY 410 AAAAGATGTTTCTCATCATCTGCAACCGTTTATGSCCAACTGTAAGAAAGATACCGTGTG 469
DB 466 AGAAGCTGTTTTCAGCAGCTCAGCCACTGTGTACGGGAACCCAGTACCTGCCCTTG 525
QY 470 AGAGG---ATTTCAGATTACAGCTATGAACTCCCTATGACGACCAAGCTTTTCCCTGG 526
DB 526 ATGAGGCCACCCACCGGTTGTTGTAACACCTTACGCGAAGTCCAAAGTCTTTCATCG 585
QY 527 AAGAAATGCGCAGATATTACAGAAAGCTGAACAGAAATGGAAGATCATATTACTGAGAT 586
DB 586 AGGAATGATCCGGACCTGTGCCAGGCACAGAAGCTTTGGAACGAGCTGCTCGCT 645
QY 587 ACTTCAATCCAGTTGGGCTCATGAAGTGGGAACTCGGTGAAGATCCCAAGGCTATCC 646
DB 646 ATTTCACCCACAGGTGCGCATGCTCTGGCTGCAATGTTGGTGAAGATCCCAAGGCTATC 705
QY 647 CAATAACCTCATGCCCTTACATTACAGCAAGTGTGTTGGAAGATTGACTGAACTCAATG 706
DB 706 CCAACAACCTCATGCCCTTATGTCTCCAGGTGGGATCGGGCAGCGGAGGCCCTGAATG 765
QY 707 TATACGTCATGATATATCCAAAGGATGGCTCTGCGATCCGGACTATATCCATGTGA 766
DB 766 TCTTTGGCAATGACTATGACACAGAGATGGCAGGTGTCCGGGATATCATCATGTCG 825
QY 767 TGGACTTGGCAGATGGCCATTTGCTCCCTGGGAAGCTCTTCAACAGGAGAACATAG 826
DB 826 TGGATCTGGCCAGGGCCCAATTGACGCTTAAAGAGCT-----GAAAGACAGTGTG 879

QY 827 GTTGTACTGCTTACAACCTGGGAACCTGCTGTGGAACATCTGTGCTTGAATGTTACAG 886
DB 880 GCTGCCGATCTACACCTGGGCAACCGGCACAGGCTATTTCAGTGTGCAGATGTTCCAGG 939
QY 887 CATTTGAAAAGGCTTTCGCAAGAAAATTCAGTAAATATATGTCAGAGAACCGGAG 946
DB 940 CTATGGAGAAGGTCTCTCTGGGAAGAATCCCGTACAAAGGTGGTGGCAGCGCGGAAGGTG 999
QY 947 ATGCGACTGAGGTTTATGTCATCTACAGAGAGCTGAGAAAGAACTTGTGGAAGGCAA 1006
DB 1000 ATGTCGAGTCTGTTAGCCCAACCCAGTCTGCCCAAGAGGAGCTGGGTGGACAGCAG 1059
QY 1007 ACTATGCTGTGGAGGAGATGTGACAGGACCAATGGAATTTGGCAAGAAACAATCCCTGGG 1066
DB 1060 CCTTAGGCTGGACAGGATGTGTAGGATCTCTGGCGCTGCGAGAGCAGATCTCTTCAG 1119
QY 1067 GTTACGCGGGAAGCCTTGA 1086
DB 1120 GCTTTGGCAGCAAGCCTGA 1139

RESULT 11

US-09-543-681A-776

; Sequence 776, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543.681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 776

; LENGTH: 1038

; TYPE: DNA

; ORGANISM: Proteus mirabilis

; US-09-543-681A-776

Query Match 17.5%; Score 229; DB 4; Length 1038;
Best Local Similarity 53.9%; Pred. No. 1.1e-55;
Matches 550; Conservative 0; Mismatches 450; Indels 21; Gaps 3;

QY 53 ACATTCTGGTCCCGGTGGTCCCGTTTCATTGGCACCCACACCGTCGTTCCAGCTTCTCA 112
DB 26 AAATTTTAGTTACCGGTGGAAATGGGATATATCGGTAGTCATACATGCGTGCATATGTTAG 85
QY 113 AAGCTGGCTTCAGCGTTTCAATAATCGACAATTTTCGATACTCCGTTCATGGAAGCAGTGG 172
DB 86 AAGCAGGATGACGCTTATTTTGGATAACCTCAGTAATGCAATGAAGAGTACTTA 145
QY 173 ACCGGTCCGCAAGTGGTGGCCCTCTGCTTTCTCAGAACCTCCAATTCACCCAGGCG 232
DB 146 ACCG-----AATTGAAGCAATTAACAGAAAACGCCACTGTTTTATCAGGGGG 193
QY 233 ATCTCCGGAATAGGATGACTTGGAGAAACTCTTCTCAAAACCAACATTTGATCCCGTGA 292
DB 194 ATATTCGTGAGAAAACAGACTGGCCCAATTTTTTACTCAGCACTCTATTCAATCAGTGA 253
QY 293 TCCACTTTGCTGGCTTGAAGCGGTGTGTAAGCGTTGCGAAGCCCGCTGCTATTTTG 352
DB 254 TCCATTTCGCTGGATTAAAGCGGTAGTGAATCTGTTGAAAAGCCCAATCGAATACTATG 313
QY 353 ATTTTAATTTGGTGGCAACCATCAACCTCTACGAGTTTATGGCAAAAGTATAATTGCAAAA 412
DB 314 ACAACAAATGTTAAACGTTACTTTGGTATTTAGTTGTTGTTATCGGTGATCGGGCGTAAAA 373
QY 413 AGATGTTTCTCATCATCTGCAACCGGTTTATGGCCAACTGAAAGATACCGTGTGAGG 472
DB 374 GTATTATTTCAGCTCATCAGCCAGGTTTATGGTGTGATCTCTCAAACTGTGCTATTACTG 433

QY 473 AGGATTTCAAGTTACAAG---CTATGATCCCTATGGACGACCAAGCTTTTCCTGGAAG 529
Db 434 AAGATCACTGTTGGTGGCACTACCAATCTTATGGCACCGTAAATATATATGTCGAGC 493
QY 530 AAATGCCCGAGATATTAGAAAGCTGAACAGAAATGAAGATCATATTAAGTGAAGTACT 589
Db 494 GTATTTTGTCTGATCTATATTCAGATGAGAGTTGGTCTATCAGTTTACTGGCTACT 553
QY 590 TCAATCCAGTTGGGCTCATGAAGTGGCAAGCTGGTGAAGATCCCAAGGGCATCCAA 649
Db 554 TCAATCCAGTAGTGTCTATCATCTGGACATATGGGGGAAGATCCTAAGGGTATTCCTA 613
QY 650 ATAACTCATGCTTACATTTACAGAAAGTAGCTGTTGGAAGATTGACTGAACCTCAATGAT 709
Db 614 ATAAATTAACGCTTATATCTCGCAAGTTGCGATAGTCTGCTGTAAGAAAGTCGGGTAT 673
QY 710 ACGGTATGATTTATCAACGAGGAGTGGCTCTGCGATCCGGGACTATATCCATGATGG 769
Db 674 TTGGTGATGATTAACCAACAAAGACGGTACTGTTGAGTACTATATCCATGTTATGG 733
QY 770 ACTTGCGAGATGGCATATTGCTGCCCTGCGAAAGCTCTTCAACGAGAGAACATAGTT 829
Db 734 ATTTAGCTATGTTATATTTGCGCTTAAATGCGTT-----AGGAGAAAAGCAGGGT 787
QY 830 GTACTGCTTACAACTGGGAATGCTGCTGGTGAACATCTGTGCTTGAATGTTACAGCAT 889
Db 788 TACATATCTATAACCTTAGCACAGGAAATGGACCAAGTGTGATTGAAATGATTGAAGCAT 847
QY 890 TTGAAAGGCTTTGCGCAAGAAATTCAGTAAATATGTCGAAGAAAGACCGGGAGATG 949
Db 848 TTGCTAAAGCAAGTGTGAAGCAATCCCTTATGAATTCAGCCACGTCGCCCGGGAGATA 907
QY 950 CGACTGAGTTTATGATCTACAGAGAGCTGAGAAAGACTTGGTTGGAAGGCAACT 1009
Db 908 TCGCAATATGTTGTTCAAGTCTTCCCAAGACGAAAGAAAGACTTACACTGGAAGCCATTC 967
QY 1010 ATGGTGTGAGGAGATGTCAGGACCAATGGAATTTGGCAAGAACCAATCCCTGGGGTT 1069
Db 968 GCTCTATTGATGATGGCAGCGCATGCTTGGCTTGGCAATTCACAAATCTTACGGTT 1027
QY 1070 A 1070
Db 1028 A 1028

RESULT 12

US-09-489-039A-6024
; Sequence 6024, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6024
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6024

Query Match 15.2%; Score 199; DB 4; Length 1092;
Best Local Similarity 52.0%; Pred. No. 4.4e-47;
Matches 532; Conservative 0; Mismatches 470; Indels 21; Gaps 3;

QY 56 TTCTGGTCAACCGGTGGTGGCGGTTTCAATGGCACCAACCGGTGCTTCAAGTCTTCAAG 115
Db 83 TTCTGGTTACAGGTGGTGGGTTTACATTGGAAGTCATACCTGCGTTCAACTGCTGCAGC 142

QY 116 CTGGCTTCAGCGTTTCAATAATTCGAACAATTTTCGATAACTCCGTCATGGAAGCAGTGGACC 175
Db 143 AGGGACACAGAGTGGTGAATCTCTCGACAATCTCTCGAAACAGCAA-----GCGCA 190
QY 176 GGGTCGCGCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCAAATTCACCCAGGGCGATC 235
Db 191 GGGTGTGCGCGTGGTGAACGTCTCGGGCGCAAGAGCCACCTTTATTGAGGCGGATA 250
QY 236 TCCGGAATAGGATGACTTTGGAGAACTCTTCTCAAAAACAAATTTGATGATCGGTGATCC 295
Db 251 TTGTAACGAAGCGCTGATGACGGAGATCTCTCAAGATCACGCCATTTAGGCGAGTATCC 310
QY 296 ACTTTCTGCTTGAAGCGGTTGCTGAAGCGTTTGGCAAGCCCGTCCGCTCTATTGATT 355
Db 311 ACTTCGCGGGCTGAAAGCGGTCGGGGAGTCCGTTCGCAAGCGCTGGAATTAATACGACA 370
QY 356 TTAATTTGGTTGSCACCATCAACCTCTACGAGTTTATGGCAAAAGTATAATTTGCAAAAAGA 415
Db 371 ATAACTCACCGGCACACTGAAATTAGTCTCTGCCATGCGCGCGCTGGCGTGAAGAACT 430
QY 416 TGGTTTCTCATCATCTGCAACGTTTATGGCCCACTGAAAGATACCGTGTGAGGAGG 475
Db 431 TTATCTTTAGCTCTCTCCGCTACCGTCTACGGCGATCAGCCGAAATTCCTGATGTGCAAA 490
QY 476 ATTTCAAGTTTACAAGCTATG---AATCCCTATGGACGACCAAGCTTTTCTTGGAAAGAA 532
Db 491 GCTTCCGACCGGCACCCCGCAAGCCCTACGGCAAAAGCAAGCTGATGTTGGAGCAGA 550
QY 533 TTGCCCCAGATATTCAGAAAGCTGAACAGAAATGGAAGATCATATTACTGAGATCTTCA 592
Db 551 TCCTGACCGACCTGCGAAGAGCCAGCCGAGTGGAGCATTTGCCCTGCTGCGTACTTCA 610
QY 593 ATCCAGTTGGGCTCATGAAAGTGGCAAACTCCGTTGAAGATCCCAAGGGCATCCCAATA 652
Db 611 ACCGTCGCGCGCCACCCCGTCGGCGCATGGGGAGACCCCGAGGGATCCCGAACA 670
QY 653 ACCTCATGCTTACATTCAGCAAGTAGTGTGGAAGATGACTGAACTCAATGTATACG 712
Db 671 ACCTGATGCCCTACATCGCCAGGTGCGCGTTTGGCGTCCGCAATCGCTGGCCATTTTTC 730
QY 713 GTCATGATTATCCAAAGAGGGATGGCTCTGCGATCCGGACTATATCCATGTGATGACT 772
Db 731 GTACGACTATCCGACCGAAGNTGGTACCGCGTGGCGATTACATCCATGATGAGACC 790
QY 773 TGGCAGATGGCCATATTGCTGCCCTCGAAAAGCTTCTTCAACGAGAAACATAGTTGTA 832
Db 791 TCGCCGATGTTCACTGTCGCGCCATGAAAACCT-----GGCTGACAAAGCCGGGTGC 844
QY 833 CTGCTTACAACTGGGAACTGGTGGTGAACATCTGCTTGAATTCGTTTACAGCATTTG 892
Db 845 ATATCTAACACCTTGGCGCGCGGTGCGGACGAGCGTCTCGACGTAGTCAATGCTTCA 904
QY 893 AAAAGGCTTCTGGCAAGAAAATTCAGTAAAAATTTATGCCAAGAACCCGGAGATCGGA 952
Db 905 GCAAGGCTCGGCAACACCCATTAACCTATTTCCGCGCGCGCGCGGATGGGACCTCC 964
QY 953 CTGAGGTTTATGATCTACAGAGAGCTGAGAAAGACTTGGTTGGAAGGCAACTATG 1012
Db 965 CGGCTTACTGGCGGATGCTGCAAGGCCGACCGCGAGCTGAACCTGGCGGTAAACGGCA 1024
QY 1013 GTGTGGAGGAGATGTGAGGAGACCAATGGAATTTGGGCAAGAAACAATCCCTGGGGTTACG 1072
Db 1025 ACCTGGAGCAATGCGCAGGACACCTGGCACTGGCAGTCCCGTCACTCCGAGGGTTATC 1084
QY 1073 CGG 1075
Db 1085 CAG 1087

RESULT 13
US-09-313-294A-7524
; Sequence 7524, Application US/09313294A

Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalugudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 7524
LENGTH: 307
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700381880H1
US-09-313-294A-7524

Query Match 12.1%; Score 158.2; DB 4; Length 307;
Best Local Similarity 70.6%; Pred. No. 1.1e-35;
Matches 211; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 640 GGATCCCAATACCTCATGCTTACATTCAGCAAGTAGCTCTTGGAAAGATTGACTGAA 699
DB 8 GGAATCCCAACATCTTATGCCCTATGTTTCAGCAAGTTCCGTTGTAGAGGCCAGCT 67
QY 700 CTCAATGTATACGCTCATGATTATTCACCAAGGAGGAGTCTCTCGCATCCGGGACTATATC 759
DB 68 CTAACAGCTTTAGGAATGACTATGCAACAAGAGATGGGACTGGGTCCGAGATTACATC 127
QY 760 CATGTGATGACTGGCAGATGCCATATTCCTGCCCTCGGAAGCTCTTCACAAAGGAG 819
DB 128 CATGTGTTGACCTTGCTGACGACATATTTGCTGATTCAGAAAGCTTTTGGAACTCT 187
QY 820 AACATAGTGTACTGCTTACAACTCGGAAGTGGTCTGTGGAACATCTGTGCTTGAATG 879
DB 188 AGCATAGGTTGAGGCTGACAACTTTGGAACCGGAGAGGTACATCTGTGCTGGAGATT 247
QY 880 GTTACAGATTGAAAGGCTCTTGGCAAGAAATTCAGTAAATATGTCACAAAG 938
DB 248 GTTAAGCATTTGAGAAGGCTTCTGGGAAGAAATACCTCTGATTTTGTGTGAAGACG 306

RESULT 14
US-09-221-017B-1093/c
Sequence 1093, Application US/09221017B
Patent No. 644799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA: PCT/AU98/01023
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 1093:
SEQUENCE CHARACTERISTICS:
LENGTH: 4597 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc feature
LOCATION: 1...4597
US-09-221-017B-1093

Query Match 11.1%; Score 146; DB 4; Length 4597;
Best Local Similarity 51.2%; Pred. No. 1.4e-31;
Matches 397; Conservative 0; Mismatches 370; Indels 9; Gaps 2;

QY 281 TTGATCCGCTGATCCACTTTGCTGGCTTGAAGCGGTTGCTGAAAGCGTTCCGAAAGCCCC 340
DB 4549 TAGCAGGGGCTCATTCATTTTGTGCAAGTAAGCCCGTAGGCGAATCCGTGCGAAGCCGC 4490
QY 341 CTCGCTATTTTGAATTTTGAATTTGTTGGCAACATCTACAGTTTATGGCAAGT 400
DB 4489 TTTTGTACTACAGNACAATATCTCTCACTGTCTGAATATCTCTGAACTGTAGGCGAT 4430
QY 401 ATAATTGCAAAAGATGGTTTTCTCATCATCTGCAACCGTTTATGGCCAACTGGAAGA 460
DB 4429 TCGGCACCTCGTGGGATCGTATTTTCTCATCTCTGATCGGTGTACGACAGCGGAGGTAT 4370
QY 461 TACCGTGTGAGGAGATTTCAAGTTAC---AAGCTATGAATCCCTTATCGACGACCAAGC 517
DB 4369 TGCCCGTAAACGAGAGGCTCCGATACAGGAAGCTCTCTCTCCCTATGCGAATACGAAGC 4310
QY 518 TTTTCTCTGGAAGAAATTCGCGAGATATTCAGAAAGCTGAACCAAGATGGAAGATCATAT 577
DB 4309 AATCAACGAGGAGATTATCCGCGATGCTATCCATGCGGAGCAGGTACAAGGCTATAC 4250
QY 578 TACTGAGATATCTTCAATCCAGTTGGGCTCATGAAGTGGCAAACTCGTGAAGATCCCA 637
DB 4249 TTCTGCGCTACTTCAATCCGATCGGGGCAACCCCGCGCATATCGGAGAGCTTCTTA 4190
QY 638 AGGCGATCCCAATAACCTCATGCTTACATTCAGCAAGTAGCTGTGGAAGATTGACTG 697
DB 4189 ACGGATGCCACAGATCTGATCCCTTATTTGACTCAACGCGACGACGATCAGACGG 4130
QY 698 AACTCAATGTATACCGGTGATGATTTCAAGAGGAGTGGCTCTCGATCCGGGACTATA 757
DB 4129 AACTGAGGCTCTTCGCGCATGACTACGACACGCCCGCGGCTTTGTATCCGGGACTATA 4070
QY 758 TCCATGTGATGACTTGGCAGATGCCATATTTGCTGCCCTGCGAAAGCTTTTCACAAAG 817
DB 4069 TCTATGTGTGATTTGGCCAAAGCGCATGTGGCTGCCATCGAACGGATGCTGATGAAG 4010
QY 818 AGAACATAGTTGTATC-----TGCTTCAACCTCGGAACTGGTCTGTGGAACATCTGTGC 871

34 ATGGTGTCTTCTCCCAACACATTTCTGGTCACCGGTGGTGGCGGTTTCATTTGGCACCAC 93
97 AGGATGGCGTTCAGGGGAAACAATTTCTGGTAACTGGAGGAGCTGGATTTCATCGGATCTCAC 156
94 ACCGTGTTTCAGCTTCTCAAGCTGCTTCAGCGTTTCAATTAATCGACAATTTCCGATAAC 153
157 ACCGTGTTTCAGCTTCTCAAGCTGCTTCAGCGTTTCAATTAATCGACAATTTCCGATAAC 216
154 TCCGTATGGAAGCAGTGGACCGCGTCCGCAAGTGGTGGCGCTCTGCTTTCTCAGAAC 213
217 TCCGTATGGAAGCAGTGGACCGCGTCCGCAAGTGGTGGCGCTCTGCTTTCTCAGAAC 276
214 CTCCAAATTCACCCAGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAA 273
277 CTCCAAATTCACCCAGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAA 336
274 ACAACATTTGATCCCGTGGATCCCACTTTGCTGGCTTGAAGCGGT 317
337 ACCAAATTTGATCCCGTGGATCCCACTTTGCTGGCTTGAAGCGGT 380

Search completed: March 22, 2004, 20:20:05
Job time : 127 secs

Qy 34 ATGGTGTCTTCTCCCAACACATTTCTGGTCACCGGTGGTGGCGGTTTCATTTGGCACCAC 93
Db 97 AGGATGGCGTTCAGGGGAAACAATTTCTGGTAACTGGAGGAGCTGGATTTCATCGGATCTCAC 156
Qy 94 ACCGTGTTTCAGCTTCTCAAGCTGCTTCAGCGTTTCAATTAATCGACAATTTCCGATAAC 153
Db 157 ACCGTGTTTCAGCTTCTCAAGCTGCTTCAGCGTTTCAATTAATCGACAATTTCCGATAAC 216
Qy 154 TCCGTATGGAAGCAGTGGACCGCGTCCGCAAGTGGTGGCGCTCTGCTTTCTCAGAAC 213
Db 217 TCCGTATGGAAGCAGTGGACCGCGTCCGCAAGTGGTGGCGCTCTGCTTTCTCAGAAC 276
Qy 214 CTCCAAATTCACCCAGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAA 273
Db 277 CTCCAAATTCACCCAGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAA 336
Qy 274 ACAACATTTGATCCCGTGGATCCCACTTTGCTGGCTTGAAGCGGT 317
Db 337 ACCAAATTTGATCCCGTGGATCCCACTTTGCTGGCTTGAAGCGGT 380

Search completed: March 22, 2004, 20:20:05
Job time : 127 secs

RESULT 15

US-09-424-311-1
; Sequence 1, Application US/09424311
; Patent No. 6372477
; GENERAL INFORMATION:
; APPLICANT: JORSBOE, Morten
; BRUNSTEDT, Janne
; PETERSEN, Steen Guldager
; TITLE OF INVENTION: CLONING OF UDP-GALACTOSE EPIMERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/424,311
; FILING DATE: 24-Jan-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/IB98/00886
; FILING DATE: 27-MAY-1998
; APPLICATION NUMBER: GB 9710991.2
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 078883/0114
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5393
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 85..381
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-424-311-1

Query Match 11.0%; Score 144.8; DB 4; Length 381;
Best Local Similarity 69.4%; Pred. No. 8.7e-32;
Matches 197; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 19:15:04 ; Search time 492 Seconds
(without alignments)
9864.409 Million cell updates/sec

Title: US-09-913-064A-13

Perfect score: 1312

Sequence: 1 gcacgagccactctctccc.....aaaaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1279.4	97.5	1481	12	US-10-424-599-36584
2	1139	86.8	1424	12	Sequence 36584, A
3	656.2	50.0	2934	12	Sequence 8664, Ap
4	644.4	49.1	1056	9	US-10-424-599-67518
5	644.4	49.1	1056	11	Sequence 67518, A
6	603.2	46.0	1488	14	Sequence 2357, Ap
7	601.8	45.9	1056	14	Sequence 2357, Ap
8	539.4	41.1	1474	12	US-10-060-275-1
9	433	33.0	1575	12	US-10-425-114-29813
10	431.4	32.9	1566	12	Sequence 9965, A
11	431.4	32.9	1622	12	Sequence 29813, A
12	430.2	32.8	1497	12	Sequence 89167, A
13	430.2	32.8	1531	12	Sequence 29766, A
14	403.8	30.8	1200	12	Sequence 5152, Ap
15	402.2	30.7	1729	12	Sequence 1181, Ap
					Sequence 22869, A
					Sequence 3333, Ap

16	395.8	30.2	1661	12	US-10-425-114-21374
17	361.2	27.5	1266	12	Sequence 21374, A
18	337.2	25.7	1850	12	Sequence 16649, A
19	312	23.8	641	9	Sequence 89166, A
20	286.8	21.9	1334	12	Sequence 615, App
21	280.5	21.4	1017	12	Sequence 3010, Ap
22	280.5	21.4	1213	12	Sequence 8209, Ap
23	278.6	21.2	2647	12	Sequence 89163, A
24	277.8	21.2	5793	10	Sequence 31639, A
25	277.8	21.2	5943	10	Sequence 2, Appli
26	272.8	20.8	1061	12	Sequence 6052, Ap
27	261.6	19.9	13206	12	Sequence 33, Appl
28	248.5	18.9	545	15	Sequence 606, App
29	247.2	18.8	3140	12	Sequence 158, App
30	244	18.6	1047	9	Sequence 9, Appli
31	244	18.6	1047	14	Sequence 17, Appli
32	244	18.6	1488	9	Sequence 97, Appl
33	244	18.6	1488	9	Sequence 547, App
34	244	18.6	1488	9	Sequence 142, App
35	244	18.6	1488	9	Sequence 360, App
36	244	18.6	1488	14	Sequence 16, Appli
37	241.4	18.4	1830121	14	Sequence 1, Appli
38	241.4	18.4	1830121	15	Sequence 1, Appli
39	240.8	18.4	1249	13	Sequence 1, Appli
40	239.2	18.2	1053	13	Sequence 64, Appl
41	231.6	17.7	874	12	Sequence 7955, Ap
42	230	17.5	1730	12	Sequence 89168, A
43	225.4	17.2	14483	14	Sequence 3, Appli
44	215.2	16.4	578	15	Sequence 600, App
45	211.8	16.1	1011	15	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-10-424-599-36584
; Sequence 36584, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 36584
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_133037C.1
US-10-424-599-36584

Query Match	97.5%	Score	1279.4	DB	12	Length	1481
Best Local Similarity	99.5%	Pred. No.	0				
Matches	1283	Conservative	0	Mismatches	6	Indels	0
				Gaps	0		
QY	8	CCACTTCTCTCCCTCTCTATTGCGAGCATGGTCTCTCTCCCAACACATCTGGTCACCG	67				
Db	111	CCACTTCTCTCCCTCTCTATTGCGAGCATGGTCTCTCTCCCAACACATCTGGTCACCG	170				
QY	68	GTGGTCCCGGTTTCATTTGGCACCACACCGTCGTTCTCAAGTGGTTCAGCG	127				
Db	171	GTGGTCCCGGTTTCATTTGGCACCACACCGTCGTTCTCAAGTGGTTCAGCG	230				
QY	128	TTTCAATATCGACATTTTCGATTAATCGTATGGAACAGTGGACCGCGTCCGCAAG	187				
Db	231	TTTCAATATCGACATTTTCGATTAATCGTATGGAACAGTGGACCGCGTCCGCAAG	290				

QY 188 TGGTTGGCCCTCTGCTTTCTCAGAACCTCGAATTACCCAGCGCGATCTCCGGATAGG 247
DB 189 TGGTTGGCCCTCTGCTTTCTCAGAACCTCGAATTACCCAGCGCGATCTCCGGATAGG 248
QY 248 ATGACTTGGAGAACTCTTCTCCAAACCAACATTTGATCCCGTGCATCTTGTGGCT 307
DB 249 ATGACTTGGAGAACTCTTCTCCAAACCAACATTTGATCCCGTGCATCTTGTGGCT 308
QY 308 TGAAGCGGTGCTGAAAGCGTTGCGAAGCCCGTGCATTTTGAATTTAAATTTGGTTG 367
DB 309 TGAAGCGGTGCTGAAAGCGTTGCGAAGCCCGTGCATTTTGAATTTAAATTTGGTTG 368
QY 368 GCACCACTCAACCTCTACGAGTTTATGCGCAAGTATAATTTGCAAAAGATGGTTTCTCAT 427
DB 369 GCACCACTCAACCTCTACGAGTTTATGCGCAAGTATAATTTGCAAAAGATGGTTTCTCAT 428
QY 428 CATCTGCAACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGAGGATTTCAAGTTAC 487
DB 429 CATCTGCAACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGAGGATTTCAAGTTAC 488
QY 488 AAGCTATGAATCCCTATGACGCGACCAAGCTTTTCTGGAAGAAATTTGCCGAGATATTC 547
DB 489 AAGCTATGAATCCCTATGACGCGACCAAGCTTTTCTGGAAGAAATTTGCCGAGATATTC 548
QY 548 AGAAGCTGAACAGAAATGGAAGATCATATTAATGAGATCTTCAATCAGTTGGGCTC 607
DB 549 AGAAGCTGAACAGAAATGGAAGATCATATTAATGAGATCTTCAATCAGTTGGGCTC 608
QY 608 ATGAAAGTGGAACTCGGTGAGATCCCAAGGGATCCCAATTAACCTCATGCCCTACA 667
DB 609 ATGAAAGTGGAACTCGGTGAGATCCCAAGGGATCCCAATTAACCTCATGCCCTACA 668
QY 668 TTGAGCAAGTAGCTGTGGAAGATTGACTGAACTCAATGTATACCGTCAATGATTCAAA 727
DB 669 TTGAGCAAGTAGCTGTGGAAGATTGACTGAACTCAATGTATACCGTCAATGATTCAAA 728
QY 728 CGAGGATGGCTCTGGATCCGGGACTATATCCATGTGATGACCTTGCAGATGGCCATA 787
DB 729 CGAGGATGGCTCTGGATCCGGGACTATATCCATGTGATGACCTTGCAGATGGCCATA 788
QY 788 TTGCTGCCCTGGAAAGCTCTTCAACCGGAGACATAGTTGTACTGTCTACAACTGG 847
DB 789 TTGCTGCCCTGGAAAGCTCTTCAACCGGAGACATAGTTGTACTGTCTACAACTGG 848
QY 848 GAACTGGTGTGGAACATCTGTGCTGAAATGGTTACAGCTTTGAAAGGCTTCTGCA 907
DB 849 GAACTGGTGTGGAACATCTGTGCTGAAATGGTTACAGCTTTGAAAGGCTTCTGCA 908
QY 908 AGAAATTCAGTAAATTTATGTCGAAGAGACCGGGAGATCGGACTGAGGTTTATGCAT 967
DB 909 AGAAATTCAGTAAATTTATGTCGAAGAGACCGGGAGATCGGACTGAGGTTTATGCAT 968
QY 968 CTACAGAGAGCTGAGAAGAACTTGTGTTGGAAGGCAAACTATGGTGTGAGAGATGT 1027
DB 969 CTACAGAGAGCTGAGAAGAACTTGTGTTGGAAGGCAAACTATGGTGTGAGAGATGT 1028
QY 1028 GCAGGACCAATGGAATTCGGCAAGAACAACTCCCTGGGTTTACCGGGAGCCCTGAA 1087
DB 1029 GCAGGACCAATGGAATTCGGCAAGAACAACTCCCTGGGTTTACCGGGAGCCCTGAA 1088
QY 1088 TTAGCTTGAGAAATATCTGCTCATCTACGAATGCTTTTCAATAAATAGGCACTCTTA 1147
DB 1089 TTAGCTTGAGAAATATCTGCTCATCTACGAATGCTTTTCAATAAATAGGCACTCTTA 1148
QY 1148 TATAGAAATCTTTTATGTTTGTATGTTTGTAGGAGTTCGTTGTATAATCTTGACAA 1207
DB 1149 TATAGAAATCTTTTATGTTTGTATGTTTGTAGGAGTTCGTTGTATAATCTTGACAA 1208
QY 1208 AAAATTTGGAGCAATTTCAAGAGTTAAAGCTATGTATTTAAACAATACTTTAAATTA 1267
DB 1209 AAAATTTGGAGCAATTTCAAGAGTTAAAGCTATGTATTTAAACAATACTTTAAATTA 1268
QY 1268 AACTTGGCCCTCTGCTTTCTCAGAACCTCGAATTACCCAGCGCGATCTCCGGATAGG 1296
DB 1269 AACTTGGCCCTCTGCTTTCTCAGAACCTCGAATTACCCAGCGCGATCTCCGGATAGG 1297

QY 1268 GACTGGCCATTGATTGATATTGAAAAA 1296
DB 1269 GACTGGCCATTGATTGATATTGAAAAA 1297
RESULT 2
US-10-425-114-8664
; Sequence 8664, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73:28
; SEQ ID NO 8664
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700787631_PLI
US-10-425-114-8664

Query Match 86.8%; Score 1139; DB 12; Length 1424;
Best Local Similarity 94.0%; Pred. No. 2.6e-300;
Matches 1195; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 26 ATTGACATGCTGTCTTCTCCCAACACATTTCTGTCACCGTGTGCGGTTTCATG 85
DB 27 ACTCCAAATGGTGACCTGTCGCAAAAGATTTCTGTTACCGCGGTGCGGTTTCATG 132
QY 86 GCACCCACACCTGCTTCAGCTTCTCAAGCTGCTTCAGCTTTCATTAATCGCAAT 145
DB 87 GCACCCACACCTGCTTCAGCTTCTCAAGCTGCTTCAGCTTTCATTAATCGCAAT 146
QY 146 TCGATACTCCGTCATGGAAGCAGTGGACCGCTCCGCCAAGTGGTGGCCCTCTGTT 205
DB 147 TCGATACTCCGTCGTTGGAAGCAGTGGACCGCTCCGCCAAGTGGTGGCCCTCTGTT 206
QY 206 CTGAAACCTCCAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAACTCT 255
DB 207 CTGAAACCTCCGAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAACTCT 256
QY 253 CTGAAACCTCCGAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAACTCT 312
DB 254 CTGAAACCTCCGAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAACTCT 313
QY 266 TCTCCAAAACAAATTTGATGCGCTGATCCACCTTCTGCTGAAAGCGGTTGCTGAA 325
DB 267 TCTCCAAAACAAATTTGATGCGCTGATCCACCTTCTGCTGAAAGCGGTTGCTGAA 326
QY 313 TCTCCAAAACAAATTTGATGCGCTGATCCACCTTCTGCTGAAAGCGGTTGCTGAA 372
DB 314 TCTCCAAAACAAATTTGATGCGCTGATCCACCTTCTGCTGAAAGCGGTTGCTGAA 373
QY 326 GCGTTGGAAGCCCGTCTGCTATTTGATTTTAAATTTGTTGGTGGCAACCTCTACG 385
DB 327 GCGTTTCCAAAGCCCGTCTGCTATTTGATTTTAAATTTGTTGGTGGCAACCTCTACG 386
QY 373 GCGTTTCCAAAGCCCGTCTGCTATTTGATTTTAAATTTGTTGGTGGCAACCTCTACG 432
DB 374 GCGTTTCCAAAGCCCGTCTGCTATTTGATTTTAAATTTGTTGGTGGCAACCTCTACG 433
QY 386 AGTTTATGCAAGATTAATTTGCAAAAGATGCTTCTCATCATCTGCAACCGTTTATG 445
DB 387 AGTTTATGCAAGATTAATTTGCAAAAGATGCTTCTCATCATCTGCAACCGTTTATG 446
QY 433 AGTTTATGCAAGATTAATTTGCAAAAGATGCTTCTCATCATCTGCAACCGTTTATG 492
DB 434 AGTTTATGCAAGATTAATTTGCAAAAGATGCTTCTCATCATCTGCAACCGTTTATG 493
QY 446 GCCAACTGAAAAGATACCGTGTGAGAGGATTTCAAGTTTCAAGTATGAATCCCTATG 505
DB 447 GCCAACTGAAAAGATACCGTGTGAGAGGATTTCAAGTTTCAAGTATGAATCCCTATG 506
QY 493 GCCAACTGAAAAGATACCGTGTGAGAGGATTTCAAGTTTCAAGTATGAATCCCTATG 552
DB 494 GCCAACTGAAAAGATACCGTGTGAGAGGATTTCAAGTTTCAAGTATGAATCCCTATG 553
QY 506 GAGGACCAAGCTTTTCTGGAAGAAATTTGCCGAGATATTCAGAAAGCTGACCAAGAT 565
DB 507 GAGGACCAAGCTTTTCTGGAAGAAATTTGCCGAGATATTCAGAAAGCTGACCAAGAT 566
QY 553 GAGGACCAAGCTTTTCTGGAAGAAATTTGCCGAGATATTCAGAAAGCTGACCAAGAT 612
DB 554 GAGGACCAAGCTTTTCTGGAAGAAATTTGCCGAGATATTCAGAAAGCTGACCAAGAT 613
QY 566 GGAAGATCATATTTACTGAGATCTTCAATCCAGTTGGGCTCATGAAAGTGGCAACTCG 625
DB 567 GGAAGATCATATTTACTGAGATCTTCAATCCAGTTGGGCTCATGAAAGTGGCAACTCG 626
QY 613 GGAAGATCATATTTACTGAGATCTTCAATCCAGTTGGGCTCATGAAAGTGGTAAACTCG 672
DB 614 GGAAGATCATATTTACTGAGATCTTCAATCCAGTTGGGCTCATGAAAGTGGTAAACTCG 673

Qy	626	GTGAAGATCCCAAGGGGCATCCCAAAATAACCTCATGCTTACATTCACGAAGTAGCTGTTG	685
Dd	673	GTGAAGATCCCAAGGGGCATCCCAAAATAACCTCATGCTTACATTCACGAAGTAGCTGTTG	732
Qy	686	GAAGATTGACTGAACCTCAATGTATACGGTTCATGATTATCCAAAGAGGGATGGCTCTGCGA	745
Dd	733	GAAGATTGCGGGAACCTCAATGTATATGSCCATGATTTATCCCAAGGGATGGCTCTGCGA	792
Qy	746	TCCGGGACTATATCCATGTGATGGACTTTGGCAGATGGCCATATTTGCTGCCCTCGGAAAGC	805
Dd	793	TCCGGGACTATATCCATGTGATGGACTTTGGCAGATGGCCATATTTGCTGCCCTCGGAAAGC	852
Qy	806	TCTTCCAAACGGAGAACATAGGTTGTACTCTCTTACCAACCTGGGAACTGGTCGTGGAAACAT	865
Dd	853	TCTTCCAAACGGAGAACATAGGTTGTACTCTCTTACCAACCTGGGAACTGGTCGTGGAAACAT	912
Qy	866	CTGTGCTGAAATGGTTTACAGCATTTGAAAGGCTTCTTGGCAAGAAAAATCCAGTAAAAAT	925
Dd	913	CCGTGCTTGAATGGTTTACAGCATTTGAAAGGCGTCTGGCAAGAAAAATCCAGTAAAAAT	972
Qy	926	TATGTCCAAAGAACCGGGAGATGCGACTGAGGTTTTATGTCATCTACAGAGAGCTGAGA	985
Dd	973	TATGTCCAAAGAACCGGGAGATGCTACTAGGTTTTATGTCATCTACAGAGAGCTGAGA	1032
Qy	986	AAGAACTTGTTTGGAAAGGCAAACTATGTGTGTGAGGAGATGTGCAGGGACCAATGGAAAT	1045
Dd	1033	AAGAACTTGTTTGGAAAGGCAAAATATGTTGTGGATGAGATGTGCAGGGACCAATGGAAAT	1092
Qy	1046	GGCCAAAGACAACTCCCTGGGTTTACGCGGGGAGGCTTGAATTAGCTTGAGAAATATAC	1105
Dd	1093	GGCCAAAGACAACTCCCGAGGTTTACGCGGGGAGGCTTGAATTAGCTTGAGAAATATAC	1152
Qy	1106	TGTCATCTACGAATGCTTTTTCACATAAATAGGCATCTCTTATATAGAATACATTTATGT	1165
Dd	1153	TGTCATCTACGAATGCTTTTCCACACAAATAGGCATCTCTT - TATAGAATACATTTATGT	1211
Qy	1166	TTGATGATTTGTTTGGCAGTTCGTTCTATAAATCTTGACAAATAAATTTGGCAGCATTT	1225
Dd	1212	TGATGATTTGTTTGGCAGTTCGTTCTATAAATCTTGACAAATAAATTTGGCAGCATTT	1271
Qy	1226	CAAGAAGTTAAAGCTATGTTATTTAAACAATAACTTTAAATTAGACTGGCCATGATTTGA	1285
Dd	1272	CAAGAAGTTAAAGCTATGTTATTTAAACAATACTTTAAATTAGACTGGCCATGATTTGA	1331
Qy	1286	TATTGAAAAAA 1296	
Dd	1332	TATTGAGAGAA 1342	

RESULT 3

US-10-424-599-67518

; Sequence 67518, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 67518

; LENGTH: 2934

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_31983C.1

US-10-424-599-67518

Db 2339 GGACAGCAAAATCCATGGGATACCAAGGGAGCATTAAATGAAATGGGTTGTA 2395

RESULT 4
US-09-938-842A-2357
; Sequence 2357, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2357
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2357

Query Match 49.1%; Score 644.4; DB 9; Length 1056;
Best Local Similarity 76.3%; Pred. No. 3.1e-165;
Matches 792; Conservative 0; Mismatches 246; Indels 0; Gaps 0;

QY 49 CAACACATCTGGTCACCGGTGGTCCGGTTTCATTGGCACCCACACCGTTCAGCTT 108
Db 19 CAGACATCTGTACTGTGGTGGTCTGCTGCTTATCGGAGCGCATCTGTGTCACTT 78
QY 109 CTCAAAGCTGGCTTCAGCGTTTCAATATCGACAAATTCGAAATTCGATCGAAGCA 169
Db 79 CTCAAAGATGGTTTTAAGGTTTCGATCATCGATAATTTTGATACTCTGTTCGAACT 138
QY 169 GTGGACCGCTCGCGCAAGTGGTGGCCCTCTGCTCTCAGAACCTCCAAATTCACCCAG 228
Db 139 GTTGATAGATTAGGAGCTTGTGGTCTGATCTCTCAGAGAGCTCGACTTCAATCTG 198
QY 229 GCGCATCTCCGAATAGGATGACTTTGAGAACTCTTCTCCAAACAACTTTTGTATGCC 288
Db 199 GGTGATCTAAGAAACAAAGGGGACATTCGAAACTATTTCTCCAGCAGAGATTTGATGCT 258
QY 289 GTGATCCACTTTGCTGGCTTGAAGCGTTGCTGAAGCGTTGCGAGCCCGCTCGCTAT 348
Db 259 GTGATTCATTTTGGCGGCTTAAAGCTGGTGGTGAAGTGTGAAACCCCTCGCGGCTAC 318
QY 349 TTTGATTTTAAATTTGGTTGGCACCATCAACCTCTACGATTTATGGCAAGTAATAATGC 408
Db 319 TTTGACAATAACTTGGTTGGACATCAATCTATATGAGACCATGGCAAGTACCACTGC 378
QY 409 AAAAGATGGTTTCTCATCTCCACCGTTTATGCCCAACCTGAAAGATACCGTGT 468
Db 379 AAAATGATGGTGTTCATCTCTCCACTGTTTATGGACCAACCTGAAAGATTTCCATGC 438
QY 469 GAGGAGGATTTCAAGTTCAAGCTATGAATCCCTATGACCGGACCAAGCTTTTCTCGAA 528
Db 439 ATGGAGACTTTGAATTAAGGCTATGAATCCTTATGCTGCTACTAGCTTCTTCTTGA 498
QY 529 GAAATTCGCGAGATATTTCAGAAAGCTGAAACAGAAATGGAAGATCATATTAAGTAC 589
Db 499 GAAATAGCTAGAGATATTCAAAGGACGAAACCGGAATGGAAATTTTCTGCTAGGTAC 558
QY 589 TTTCAATCCAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCA 648

Db 559 TTCAATCCTGTAGGAGCACATGAGAGTGCGAGTATTGGTGGAGATCCAAAGGCGATCCCC 618
QY 649 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTGGAAAGATTGACTGAATCAATGTA 708
Db 619 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTGGAAAGATTGACTGAATCAATGTC 678
QY 709 TACGGTCAATGATTATCCAAAGGAGATGGCTCTGGATCCGGGACTATATCCATGTGATG 768
Db 679 TATGGACATGACTATCCCAACCGAGGATGGTAGTGGGTAAGAGACTTACATCCATGTGATG 738
QY 769 GACTTGGCAGATGGCCATATTGCTGCCCTGCGAAAGCTCTTCACAAAGGAGACATAGT 828
Db 739 GATTTAGCAGATGGCCATATCGCTGGCTCAGGAAGCTATTTCGCTGATCCAAAGATTGGT 798
QY 829 TGTACTGCTTACAACTGGGAACTGGTCTGCGAAATCTGTCTTCAAAATGGTTACAGCA 888
Db 799 TGTACTGCTTACAACTAGGACTGGTCAAGGAAAGCTGTGTGTAGAAATGGTTCAGCT 858
QY 889 TTTGAAAAGGCTTCTGGCAAGAAATTCAGTAAATATATGTCCAAAGAGACCGGAGAT 948
Db 859 TTTGAAAAGGCTTCCGGCAAGAAATTCAGTAAATATATGTCCAAAGAGGTCAGGAGAT 918
QY 949 GCGACTGAGGTTTATGCTCTACAGAGAGCTGAGAAAGAACTTGGTTGGAGGCAAAAC 1008
Db 919 GCAACAGCAGTTTATGCTTCAACAGAGAGGCTGAGAAAGAACTTGGCTGGAGGCAAA 978
QY 1009 TATGTTGTGGAGAGATGTGAGGAGCAATCGAAATTTGGGCAAGAACTCCCTGGGT 1068
Db 979 TATGGAGTGGATGAGATGTGAGAGATCAGTGGAAATGGGCAACAAATATCCATGGGT 1038
QY 1069 TACGCGGGAAGCTTGA 1086
Db 1039 TACCAGATAAGCTTTGA 1056

RESULT 5
US-09-938-842A-2357
; Sequence 2357, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2357
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2357

Query Match 49.1%; Score 644.4; DB 11; Length 1056;
Best Local Similarity 76.3%; Pred. No. 3.1e-165;
Matches 792; Conservative 0; Mismatches 246; Indels 0; Gaps 0;

QY 49 CAACACATCTGGTCACCGGTGGTCCGGTTTCATTGGCACCCACACCGTTCAGCTT 108
Db 19 CAGACATCTGTACTGTGGTGGTCTGCTGCTTATCGGAGCGCATCTGTGTCACTT 78
QY 109 CTCAAAGCTGGCTTCAGCGTTTCAATATCGACAAATTTGAAATTCGATCGAAGCA 168
Db 109 CTCAAAGATGGTTTTAAGGTTTCGATCATCGATAATTTTGATACTCTGTTCGAACT 138
QY 169 GTGGACCGCTCGCGCAAGTGGTGGCCCTCTGCTCTCAGAACCTCCAAATTCACCCAG 228
Db 139 GTTGATAGATTAGGAGCTTGTGGTCTGATCTCTCAGAGAGCTCGACTTCAATCTG 198
QY 229 GCGCATCTCCGAATAGGATGACTTTGAGAACTCTTCTCCAAACAACTTTTGTATGCC 288
Db 199 GGTGATCTAAGAAACAAAGGGGACATTCGAAACTATTTCTCCAGCAGAGATTTGATGCT 258
QY 289 GTGATCCACTTTGCTGGCTTGAAGCGTTGCTGAAGCGTTGCGAGCCCGCTCGCTAT 348
Db 259 GTGATTCATTTTGGCGGCTTAAAGCTGGTGGTGAAGTGTGAAACCCCTCGCGGCTAC 318
QY 349 TTTGATTTTAAATTTGGTTGGCACCATCAACCTCTACGATTTATGGCAAGTAATAATGC 408
Db 319 TTTGACAATAACTTGGTTGGACATCAATCTATATGAGACCATGGCAAGTACCACTGC 378
QY 409 AAAAGATGGTTTCTCATCTCCACCGTTTATGCCCAACCTGAAAGATACCGTGT 468
Db 379 AAAATGATGGTGTTCATCTCTCCACTGTTTATGGACCAACCTGAAAGATTTCCATGC 438
QY 469 GAGGAGGATTTCAAGTTCAAGCTATGAATCCCTATGACCGGACCAAGCTTTTCTCGAA 528
Db 439 ATGGAGACTTTGAATTAAGGCTATGAATCCTTATGCTGCTACTAGCTTCTTCTTGA 498
QY 529 GAAATTCGCGAGATATTTCAGAAAGCTGAAACAGAAATGGAAGATCATATTAAGTAC 589
Db 499 GAAATAGCTAGAGATATTCAAAGGACGAAACCGGAATGGAAATTTTCTGCTAGGTAC 558
QY 589 TTTCAATCCAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCA 648

QY 830 GTACTGCTTACACCTGGGAAGTGGTGGGAACATCTGTGCTTGAATGGTTACAGCAT 889
Db 942 GTGTTGCTCACTAATTTGGGTACTGGAAAGCAAAATCTGCTTAGAGATGGTCTGCTT 1001
QY 890 TTGAAGAAGCTTCTGGCAAGAAATTCAGTAAATATATGTCGAAGAGACCGGGAGATG 949
Db 1002 TTGAAGAAGCTTCTGGCAAGAAATTCAGTAAATATATGTCGAAGAGACCGGGAGATG 1061
QY 950 CCACTGAGTTTATCATCTACAGAGAGCTGAGAAAGAACTTGGTTGGAGCAAACT 1009
Db 1062 CCACTGAGTTTATCATCTACAGAGAGCTGAGAAAGAACTTGGTTGGAGCAAACT 1121
QY 1010 ATGGTGTGAGAGATGTCAGGACCAATGGAATGGGCAAGAACTTGGTTGGAGCAAACT 1069
Db 1122 ATGGTGTGAGAGATGTCAGGACCAATGGAATGGGCAAGAACTTGGTTGGAGCAAACT 1181
QY 1070 ACGCGGGAAGCTTGAATAGCTTGAAGAAATATAC 1105
Db 1182 ACCAATCAAGGCTTGAATAGCTTGAAGAAATATAC 1217
RESULT 7
US-10-060-275-3/c
; Sequence 3, Application US/10060275
; Publication No. US20030073828A1
; GENERAL INFORMATION:
; APPLICANT: DAI, ZIYU
; APPLICANT: SHI, LIFANG
; APPLICANT: HOOKER, BRIAN S.
; TITLE OF INVENTION: EPIMERASE GENE AND USE THEREOF
; FILE REFERENCE: 059440-0143
; CURRENT APPLICATION NUMBER: US/10/060,275
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/265,311
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Solanum tuberosum
US-10-060-275-3
Query Match 45.9%; Score 601.8; DB 14; Length 1056;
Best Local Similarity 73.8%; Pred. No. 1.4e-153;
Matches 765; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
QY 50 AACACATTCTGGTCACCGGTGGTGGCGGTTTCATTGGCAACCCACACCGCTGTTCAAGCTTC 109
Db 1037 AAAATATTTGGTTACTGAGGAGCTGGTTTCATTGGACACACACTGTGGTCAAGTTAC 978
QY 110 TCAAGCTGGTTACAGCTTTCATTAATCGACAATTCGATTAACCTCCGTCATGGAAGCAG 169
Db 977 TGAATGAAGGGTTCAAAGTTACCATCAATTGATAACTTTTCATAATCTGTGGAAGAAGCTG 918
QY 170 TGGACCGCTGCCCAAGTGGTTGGCCCTCTGCTTCTCAGAACCTCCAAATTCACCCAGG 229
Db 917 TTGATAGATCAGAGAAATAGTTGGTCTCACTTTCAGAAATCTTGAATTCATTTGG 858
QY 230 GCGATCCCGAATAGGATAGCTTGGAGAACTCTTCCAAAACAACTTGAATGACCG 289
Db 857 GTGATATTGAACAAGAAGATGACTTGGAGAGCTATTTCTAAGAAGAAGATTGCTGCTG 798
QY 290 TGTATCCACTTGGCTGGCTTGAAGCGTTGCTGGAAGCGTTGGAAGCCCGCTGCTATT 349
Db 797 TGTGTCATTTGCTGGCACTTAGCTGTTGGAGAGAGTGTGTTGAGCCCTTCTTACT 738
QY 350 TTGATTTTAAATTTGGTTGGCAACATCAACTCTACAGTTTATGGCAAAAGTAAATGCA 409
Db 737 TTGAGAACATCTGATTGGATCAATAACTTTGATTTAGTCATGCGCAAGTAAATGTA 678
QY 410 AAAAGATGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTG 469

Db 677 AGAAGTTGGTTTTTTCATCATCTGACAGTTTATGTCAGCTGAAAAGTTCCCTGTG 618
QY 470 AGAGATTTCAAGTTTCAAGCTATGAATCCCTATGAGCGGACCAAGCTTTTCTTGGAG 529
Db 617 TGGAGATTTTCAATTTGAAGGCTATGAATCCCTATGAGCGGACCAAGCTTTTCTTGGAG 558
QY 530 AAATTCGCCGAGATATTGAGAAAGCTGACCAAGATGGAAGATCATATTACTTGAGATACT 589
Db 557 ATATTGCTCGGATATCCAGAGGAGATCAAGATGGAATATCATCTGTTGAGTATT 498
QY 590 TCAATCCAGTTGGGCTCATGAAAGTGGCAAACTGGTGAAGATCCCAAGGCGATCCCA 649
Db 497 TCAACCCAGTAGAGCTCATGAAAGTGGCAAACTGGGGAAGATCCAAAGGCGATCCCA 438
QY 650 ATAACTCATGCTTACATTGACCAAGTAGCTGTTGGAGATTTGACTGAATGAT 709
Db 437 ACATCTTATGCTTACATTGACCAAGTAGCTGTTGGTAGATTGCCAGAGTTGAATGAT 378
QY 710 ACGTTCATGATATCCACGAGGAGTGGCTTGGGATCCGGGACTATATCCATGTGATCG 769
Db 377 ATGGCAACGACTACCTACACCTGATGTTACCGGATACGAGATTATATCCATGTTTGG 318
QY 770 ACTTGGCAGATGSCCATATTGCTGCCCTCGGAAAGCTCTTCAACCGGAGAACATAGGTT 829
Db 317 ATTTAGCGAGCGTCAATGTTTGGCACTTCAGAGACTTCTAAGGCAGATCATTTAGGTT 258
QY 830 GTACTGCTTACAACTGGGAACCTGGTGGCAACATCTGTGCTTGAATGGTTACAGCAT 889
Db 257 GTGTGCTTACAAATTTGGGTACTTGGAAAGGCAAACTCTGCTAGAGATGGTTGCTGCT 198
QY 890 TTGAAAAGCTTCTGGCAAGAAATTCAGTAAATATGTCGAAGACCGGAGATG 949
Db 197 TTGAAAAGCGCTCTGGAAGAAATATCCGCTTAAATGTTCCNAGAGACCGAGAGATG 138
QY 950 CGACTGAGGTTTATGCTCTACAGAGAGCTGAGAAAGAACTTGGTTGGAGGCAAACT 1009
Db 137 CCACTGCTGTTTATGCTACTTGAAGAGCTGAGAAAGAGCTCGGTTGGAAGCAAAAT 78
QY 1010 ATGTTGGAGAGATGTCAGGAGCAATGGAATGGCAAGAAACAACTCCCTGGGTT 1069
Db 77 ATGTTAAATGAGATGTCAGGAGCAAGTGAATGGCAAGCAAAATCCCTGGGTT 18
QY 1070 ACGCGGGAAGCTTGA 1086
Db 17 ACCAATCAAGCTTGA 1
RESULT 8
US-10-425-114-9965
; Sequence 9965, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9965
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700889134_FLI
US-10-425-114-9965

Query Match 41.1%; Score 539.4; DB 12; Length 1474;
Best Local Similarity 67.1%; Pred. No. 1.9e-136; Indels 103; Gaps 3;
Matches 865; Conservative 0; Mismatches 321;

QY 29 GCAGCATGTGTCTTCTCCCAACACATCTCTGTCACCGGTGTCGCGGTTTCATTCGCA 88
DB 89 GGAGGATGTGTCCCGATGTCAACGGTTCTGTGACGGGAGGGGAGGCTTCATCGGAT 148
QY 89 CCACACCGCTGTTCAGCTTCTCAAGCTGCTTCAGCGTTTCAATAATCGACAATTCG 148
DB 149 CCACACCGGTGTTCAGCTTCTCAAGCGGCTTCAGGGTTTCCATATCGACAACCTTG 208
QY 149 ATAACTCCGTTCATGGAAGCAGTGGACCGGTTCGCGCAAGTGGTGGCCCTCTGCTTTC 208
DB 209 ACATTTCCCTCATCGAAGCGTTTCAGAGTTTCGCGCGCTGTGCGTCTCATCTTCCA 268
QY 209 AGAAGCTCCAAATTCACCGAGGCGATCTCGGATAGGATGACTTGGAGAACTTCT 268
DB 269 ACAAGCTCACTTCTTCCACGGAGTCTCCGCAATGTCAGGATTTGGAGGCGAGTGTCT 328
QY 269 CCAAAACAACA-----TTTGATGCGGTGATCCACTTTGCTGCTTGAAGCGG 316
DB 329 CGAAGCAAACTGAACAGGTTTGTGCGGTGATCCACTTTGCTGGCTGAAAGGTG 388
QY 317 TTGCTGAAGCGTTGGGAAGCCCGTGGCTATTTGATTTTAAATTTGGTTGGCACCATCA 376
DB 389 TCGCGAAGCGTTGCAAGCCCGCGTTTATACGACCAATCTAGTGGGACTATAA 448
QY 377 ACCTTACGAGTTTATGGCAAGTATAATTCGAAAGATGGTTTCTCATCATCTGCAA 436
DB 449 ACCTTTTGAAGCAATGGCTTAATACAGTGTAAAAAATGGTTATATCATCATCGGCAA 508
QY 437 CCGTTTATGCGCAACCTGAAAGATACCGGTG---GAGGAGGATTTCAAGTACAGTAA 493
DB 509 CTGTTATGGGAAGCTGATAGAGTCCATGTGTGGAGGAGGAGTGCATTTGCGAGGCA 568
QY 494 TGAATCCCTATGACGACCAAGCTTTCTGTAAGAAATGCGCGAGATATTCAGAAAG 553
DB 569 TGAATCCGTAATGAAGAACAAAGCTGTCTGTAAGAAATGAGCCAGAGATCCAGAGG 528
QY 554 CTGAACCAAGATGAAGATCATATTTACTGAGATACCTTCAATCCAGTTGGGCTCATGAA 613
DB 529 CGGAGACAGATGAGGATCATTTCTGCTTCAATCTCAATCCGTTGGTCCCAAGAG 688
QY 614 GTGCAAACTCGGTGAAGATCCGAGGCAATCCAAATACCTTCAGCTTACATTCAGC 673
DB 689 GTGACAGATTTGGGGAAGATCCAGGGGAATCCCAATACCTTCATGCTTACATTCACC 748
QY 674 AAGTAGCTGTGGAAGATTCGACTGAATCAATGTATACGCTCATGATTATCCAAAGAGG 733
DB 749 AAGTCGCGTTGGCAGATTCGCTCAGCTCAATGTTTATGTCATGACTATCCCACTAAAG 808
QY 734 ATGGCTCTGC-----743
DB 809 ATGGCACCCCGGTCACTCACTCACTCTTGTCTTTTCATTTCTTTTCTTCTT 868
QY 744 -----GATCGGAGCTATATCCATGTG 765
DB 869 CTTTCTCTTCTTTCAGCTTACAAATATATCTGAGATTCGGGACTATATCCATGTA 928
QY 766 ATGAGCTTGGCAGATGCCATATTCGCTGCGCTGCGAAGCTTTCACAAAGGAGACATA 825
DB 929 ATGAGCTTGGCAGATGCCATATTCGCTGCGCTGCGAAGCTTTTGCACAGACCAATC 988
QY 826 GGTGTACTGTTACAACTGGGAACTGCTGCTGGAGACATCTGTGCTTGAATGGTTTACA 885
DB 989 GGTGTAGTGCTTCAATCTTTGGAACTGGGCGTGGCAGATCCGCTTGTGAATGGTTGCT 1048
QY 886 GCATTTGAAAAGGCTTCTGGCAAGAAATTCAGTAAATATATGTCACAAAGAACCCGGGA 945
DB 1049 GCTTTTGAAGGCTTCGGCAAGAAATTCATTAATAATGTCCTCCAGAAAGCCCGGG 1108
QY 946 GATGGACTGAGGTTTATGATCTACAGAGAGCTGAGAAAGAACTTGGTTGGAGGCA 1005

DB 1109 GATCTACTGCTGTATATGCTATACGGAACAAGCTGAGAAAGAACTTGGTTGGAAGGCA 1168
QY 1006 AACTATGCTGTGAGGAGATGTCAGGACCAATGGAATTTGGCAAGAAACAATCCCTGG 1065
DB 1169 AAATACGGTATAGAGGAATGTGCGAGGATTTGTGGAATTTGGCAACAAAATCCATGG 1228
QY 1066 GTTTACGCGGGAAGCCCTTGAATTAGCTTGAGAAATATCTGCTCATCTACGAATGCTTT 1125
DB 1229 GGATACAGGAGGACATTAATGAATTTGGTTGTACTTTGGAGGAGTCTTCTCCACC 1288
QY 1126 TCACATAAATAGGATCTCTTATATAGAACTTTATGTTTATGATGATTTGTTTAGGCAG 1185
DB 1289 ATATCATATCATCTCCACTCCAGATATAATATGTGATGTTCTTTTGTGATACAAATTG 1348
QY 1186 TTCGTTGTATAATCTTCACAAATAAAATTT 1214
DB 1349 ATTGATTCAATCTTTACACTACGTATGTT 1377

RESULT 9
US-10-425-114-29813
; Sequence 29813, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-211531313B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29813
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMROPIC090E09_FLI
US-10-425-114-29813

Query Match 33.0%; Score 433; DB 12; Length 1575;
Best Local Similarity 65.5%; Pred. No. 2.4e-107;
Matches 666; Conservative 0; Mismatches 345; Indels 6; Gaps 2;

QY 56 TTCTGCTCACCGGTGGTCCCGGTTTCATTTGGCACCCACACCGTTCGTTGAGTTCTTCAAG 115
DB 194 TACTGTTAAACCGGGGAGCGGTTACATCGGCAGCCACACCGTTCTTCAGCTCTTGTCTCG 253
QY 116 CTGGCTTCAGCGTTTCAATATACGACATTTTCGATACCTCCGTCATGGAAGCAGTGGACC 175
DB 254 GAGGTTTCAGAGCGCGTCTGCTCGACCAACCTCGAAAATTCCTCGAGTTGCCATGCA 313
QY 176 GGTCCCGCAAGTGGTTGGCCCTCTGCTTCTCGAAGCTCCAAATTCACCGAGCGGATC 235
DB 314 GAGTTCAGGAGCTCCCGGC---GAAATTTGGGAACAACCTCTCTCTTCAAGTGGAGC 370
QY 236 TCCGAAATAGGATGACTTGGAGAACTTCTTCTCCAAACCAACATTTGATCCGCTGATCC 295
DB 371 TACGGGACAGAGCTCTCTAGACCAAAATATTTCTTCCACAAATTCGATGCTGTCTATAC 430
QY 296 ACTTTTGTGCTTGAAGCGTTGCTGAAAGCGTTTGGCAAGCCCGCTCGCTATTTTGTATT 355
DB 431 ATTTTGTGCTGAAAGCGTGAAGAGTGAAGAGTGTGCAAAAACCTTTTACTACTATAACA 490
QY 356 TTAATTTGTTGGCACCATCAACCTCTACGAGTTTATGCGCAAGTATATTAATTGCAAAAGA 415
DB 491 ACAACTTGACTGGGCAATCACTCTATTTGGAAGTCAATGCTGCCCTGCCATGATGCAAGAAGC 550

```
QY 416 TGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGGAAGATACCGTGTGAGGAG 475
Db 551 TCGTGTCTTTCATCTTCAGCAACTGTATATGTTGGCCAAAGAGGTTCCATGCACAGAAG 610
QY 476 ATTTCAGTTTACAAGCTATGAATCCCTATGACGACCAAGCTTTTCTGGAAGAAATTG 535
Db 611 AGTTCCCTCTGTGACCAATGAACCCATATGAGCAACTAAGCTTATCATTTGAAGAAATTT 670
QY 536 CCGGAGATATTGAGAAGCTGAACAGATGAAGATCATATTACTGAGATCTTCAATC 595
Db 671 GCGGTGATGTCACCTGTGAGAGCCAGATTTGAAAATTAATTTGTTAAGATCTTCAACC 730
QY 596 CAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAATAACC 655
Db 731 CAGTTGTCACACACCCAGTGGTTATTTGGGAGGATCTCGTGGAAATCCAAACATC 790
QY 656 TCATGCTTACATTCAGCAAGTAGCTGTTGGAAAGATTGACTGAATCATATGATACGTC 715
Db 791 TCATGCTTACATTCAGCAAGTAGCTGTTGGAAAGATTGACTGAATCATATGATACGTC 850
QY 716 ATGATTATCAACAGAGGATGGCTCTGCGATCCGGGACTATATCCATGTGATGACTGG 775
Db 851 ATGATTATCAACAGAGGATGGCTCTGCGATCCGGGACTATATCCATGTGATGACTGG 910
QY 776 CAGATGCCATATTGTCGCCCTGCGAAAGCTCTTCAACAGGAGAACATAGTTTACTG 835
Db 911 CAGATGCCATATTGTCGCCCTGCGAAAGCTCTTCAACAGGAGAACATAGTTTACTG 967
QY 836 CTTACAACTGGGAACCTGCTGGGAACATCTGCTTGAATGTTACAGCATTTGAAA 895
Db 968 TTTATNACTGGGAACAGGAAGGGAACATCACTTTTGGAGATGTTAGAGCTTTGAAA 1027
QY 896 AGGCTTCTGGCAAGAAATTCAGTAAATATGTTCCAAAGAACCCGGAGATGCGACTG 955
Db 1028 TGGCATCTGGAAGAAATTCAGTAAATATGTTCCAAAGAACCCGGAGATGCGACTG 1087
QY 956 AGGTTTATCATCTACAGAGAGCTGAGAAAGACTTGGTTGGAAGCAACTATGCTG 1015
Db 1088 TTGTTTATCATCTACAGAGAGCTGAGAAAGACTTGGTTGGAAGCAACTATGCTG 1147
QY 1016 TGGAGGAGATGTGCGAGGACCAATGGAATTTGGCCAAAGAACCAATCCCTGGGTTACG 1072
Db 1148 TTGATGAGATGTGCGGCTGATCAATGGAATTTGGCTAGCAAAACCCCTTATGGCTATG 1204

RESULT 10
US-10-424-599-89167
; Sequence 89167, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David X
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 89167
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MBT3847_51526C.1
US-10-424-599-89167

Query Match 32.9%; Score 431.4; DB 12; Length 1566;
Best Local Similarity 65.4%; Pred. No. 6.6e-107;
Matches 665; Conservative 0; Mismatches 346; Indels 6; Gaps 2;
QY 56 TTCTGGTCACCGGTGGTGGCCGGTTTCATTGTCGACCCACACCGTCGCTTCAGCTTCTCAAG 115
```

RESULT 11
US-10-425-114-29766
; Sequence 29766, Application US/10425114
; Publication No. US20040034898A1
; GENERAL INFORMATION:

```
Db 185 TACTGGTAACCGGAGCGGTTTATCGGACGCCACCGTTCTTCTAGCTTCTGCTCG 244
QY 116 CTGCTTTCAGCGTTTCAATTAATCGACAAATTCGATTAACCTCGCTCATGGAAGCAGTGGACC 175
Db 245 GAGGTTTCAGAGCGGCTGCTCGACAACTCGAAAATTCCTCGAGGTTGCCATCCACA 304
QY 176 GCGTCCGCAAGTGGTTGGCCCTCTGCTTCTCAGAACCTCCCAATTCACCCAGGGCGATC 235
Db 305 GAGTCAGGAGCTGCGCGGC---GAATTTGGGAACAACTCTCTCTTTCACAGGTGGACC 361
QY 236 TCCGGAATAGGATGACTTTGGGAAGAACTCTTCTCAGAAACAAATTTGATGCGCTGATCC 295
Db 362 TACGGGACAGAGCTGCTTAGACCAATATTTCTTCCACAAATTCGATGCTGTCTATAC 421
QY 296 ACTTTCGCTGCTTGAAGCGTTGCTTGAAGCGTTGGGAGCCCGCTCGCTATTTTGAAT 355
Db 422 ATTTTGTGTCAGTGAAGCAGTAGGAGAAAGTGTGCAAAACCTTACTATATCTATAACA 481
QY 356 TTAATTTGTTGGTGGCACCATCAACTCTAGAGTTTATGGCAAGATATAATTCGCAAAAGA 415
Db 482 ACAACTTACTGGGACCAATCACTCTATTTGGAAGTCATGGCTGCCATGGATGCAAGAAGC 541
QY 416 TGGTTTCTCATCTATGCAACCGTTTATGGCCAACTGGAAGATACCGTGTGAGGAG 475
Db 542 TCGTGTCTTCTCATCTTTCAGCAACTGTATATGTTGGCCAAAGGAGTTTCCATGACAGAAG 601
QY 476 ATTTCAAGTTTACAGCTATGAATCCCTATGCGAGCGCAAGCTTTTCTCGGAAGAAATTG 535
Db 602 AGTTCCCTCTGTGACCAATGAAACCCATATGGACCACTAAGCTTATCAATTGAAGAAATTT 661
QY 536 CCGGAGATATTGAGAAGCTGAACAGATGGAAGATCATATTACTGAGATCTTCAATC 595
Db 662 GCCGTGATGTTCCACTGTGACAGCCAGATTTGTAATAATAATTTTGTAAAGATATCTCAACC 721
QY 596 CAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGCGATCCCAATAAACC 655
Db 722 CAGTTGGTGCACACCCAGCTGGTTATTTGGGGAGGATCTCTGCGGAATTCCAACAATC 781
QY 656 TCATGCTTACATTCAGCAAGTAGCTGTTGGAAAGATTGACTGAATCAATGATACGGTC 715
Db 782 TCATGCTTACATTCAGCAAGTAGCTGTTGGCGAGCGGCTGCACTGACAGTTTTCGAA 841
QY 716 ATGATTATCCACGAGGATGGCTTGGATCGGAGCTATATCCATGTGATGAGACTGG 775
Db 842 ATGATTATTAACAAGTATGGCACTGGGGTTCGGGATTAACATTCATGTTGTTGATTAG 901
QY 776 CAGATGCCATATTGTCGCCCTGCGAAAGCTCTTCAACGGGAGAACATAGTTTACTG 835
Db 902 CAGATGGGCACATTGCTGCTTAAACTAGATGAACC---TAATATAGGTTGTGAGG 958
QY 836 CTTTACAACTGGCAACTGTTGCGGAACATCTGCTTGAATGTTTACAGCAATTTGAAA 895
Db 959 TTTATTAACCTGGGAACAGGAAGGGAACATCAGTTTGGAGATGGTTAGAGCTTTTGAAA 1018
QY 896 AGGCTTCTGGCAAGAAATTCAGTAAATATGTTCCAAAGAACCGGAGATGCGACTG 955
Db 1019 TGGCATCTGGAAGAAATTTCCACTTGTGATGGCTGGCCGTAGACCTGGTGTGATGCTGAAA 1078
QY 956 AGGTTTATGCTATCAGAGAGAGCTGAGAAAGAACTTGGTTGGAAGCAACTATGCTG 1015
Db 1079 TTGTTTATGCTATCAACAAAGAAAGCGGAAGAGAGCTTAAATGGAAGGCAAAATATGGCA 1138
QY 1016 TGGAGGAGATGTGACGAGGACCAATGGAATTTGGGCAAGAACCAATCCCTGGGTTACG 1072
Db 1139 TTGATGAGATGTGCGGCTGATCAATGGAATTTGGGCTAGCAAAACCCCTTATGGCTATG 1195
```


APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 29766
LENGTH: 1622
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: UC-GWNOIR018B10_FLI
US-10-425-114-29766

Query Match 32.9%; Score 431.4; DB 12; Length 1622;
Best Local Similarity 65.4%; Pred. No. 6.8e-107;
Matches 665; Conservative 0; Mismatches 346; Indels 6; Gaps 2;

QY 56 TTCTGTCACCGTGTGCGGTTTCATTTGGCACCCACACCGTCGTTGAGCTTCTCAAG 115
DB 195 TACTGGTAACCGCGAGCGGTTACATCGGCAGCCACACCGTCTTCTGAGCTTCTGCTCG 254
QY 116 CTGCTTTCAGCGTTCAATTAATCGACAATTCGATACTCCGTCATGGAAGCAGTGGACC 175
DB 255 GAGGTTTCAGAGCGGTCCTCGACACCTCGAATTCCTCCGAGGTTGCCATCCACA 314
QY 176 GCGTCGCGCAAGTGTGGCCCTCTGCTTCTCAGAACTCCAAATTCACCCAGGGGATC 235
DB 315 GAGTCAGGAGCTCGCGGC---GAATTTGGGAACAACCTCTCTTTCACAAAGGTGGACC 371
QY 236 TCCGGAATAGGATGACTTCGGAGAACTCTTCCAAAACAACATTTGATGCGGTGATCC 295
DB 372 TAGGGACAGAGCTGCTTAGACCAATATTTCTCCACAAATTCGATGCTGCTCATC 431
QY 296 ACTTTGCTGCTTGAAGCGGTTGCTGAAAGCGTTGCGAAGCGCCGCTGCTATTGTAT 355
DB 432 ATTTGCTGACTGAAAGCAGTAGAGAGAGAGTGTGCAAAACCTTCTTACTATATAACA 491
QY 356 TTAATTTGTTGGCACCACCACTCTACGAGTTATGCAAGTATATTCGAAAGATATTCGAAAGA 415
DB 492 ACACTTGACTGGACAATCACTCTATTTGGAATCATGGCTGCGCCATGATGCAAGAAGC 551
QY 416 TGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGGAGG 475
DB 552 TCGTGTCTTCTATCTTCAGCACTGTATATGTTGGCCAAAGAGGTTCCATGCACAGAG 611
QY 476 ATTTCAAGTTACAGCTATGAATCCCTATGGACGGACCAAGCTTTTCTTGGAGAAATG 535
DB 612 AGTTCCCTCTGTGAGCAATGAACCCATATGAGCACTAAGCTTATCATTTGAAGAAATTT 671
QY 536 CCGGAGATATTCAGAAAGCTGAAACAGAGATGGAAGATCATATTACTGAGATACTTCAATC 595
DB 672 GCGGTGATGTCCTGTGACAGCCAGATTTGAAATATTTTGTAAATACTTCAACC 731
QY 596 CAGTTGGGCTCATGAAAGTGGCAACTCGGTGAAGATCCCAAGGCGATCCCAATAAAC 655
DB 732 CAGTTGGTGACACCCCGAGTGTATATTTGGGGAGGATCTCTGTTGAATTCCAACAATC 791
QY 656 TCATGCTTACATTCAGCAGTAGCTGTTGGAAGATTGACTGAATCAATGATATACGGTC 715
DB 792 TCATGCCATTTGTTACGACAGTAGCAGTTGGCCGACGGCTGCATGACAGTTTGGNA 851
QY 716 ATGATTATCAACAGGGATGCTTGGCATCCGGAGCTATATCCATGATGATGGACTTGG 775
DB 852 ATGATTATATAAAGTATGAGTGGCACTGGGGTTCGGGATTTACATTCATGTTGTTGATTAG 911
QY 776 CAGATGGCCATATTTGTCGCCCTCGGAAGCTCTTTCACACGGGAGACATAGGTTGCTACTG 835

DB 912 CAGATGGGCACATTTGCTGCAATTTGCTTAACTAGATGAACC---TAATATAGTTGTGAGG 968
QY 836 CTTACAACTGGGAACCTGGTGTGGAACAATCTGTCTTGAATGTTTACAGCATTTGAAA 895
DB 969 TTTATAACCTGGGAACAGGAAGGAAACATCAGTTTTTGGAGATGTTAGAGCTTTTGA 1028
QY 896 AGGCTTCTGGCAAGAAATTCAGTAAATTTATGTCCTCAAGAGACCGGAGATCGACTG 955
DB 1029 TGGCATCTGGAAGAAATTCACCTTGTGATGGCTGCGCTAGACCTGTTGATGCTGAAA 1088
QY 956 AGGTTTATGATCTACAGAGAGACTGAGAAAGAACTTTGTTGGAAGCAAACTATGTTG 1015
DB 1089 TTGTTTATGATCAACAAAGAAAGCGGAAGAGAGCTTTAAATGGAAGCAAAATATGCA 1148
QY 1016 TGGAGGAGATGTCAGGAGCAATGGAATGGGCAAGAAACAATCCCTGGGGTTACG 1072
DB 1149 TTGATGAGATGTCGCTGATCATTTGGAATGGGCTAGCAAAACCTTATGGCTATG 1205

RESULT 12

US-10-425-114-5152
Sequence 5152, Application US/10425114
Publication No. US2004003488A1
GENERAL INFORMATION:

APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 5152
LENGTH: 1497
TYPE: DNA
ORGANISM: Zea mays
FEATURE:

OTHER INFORMATION: Clone ID: 700445105_FLI
US-10-425-114-5152

Query Match 32.8%; Score 430.2; DB 12; Length 1497;
Best Local Similarity 65.0%; Pred. No. 1.4e-106;
Matches 668; Conservative 0; Mismatches 353; Indels 6; Gaps 2;

QY 54 CATCTGTCTACCGTGTGCGGTTTCATTTGGCACCCACACCGTCGTTGAGCTTCTCAA 113
DB 171 CATCTGTGACGGCGGCGCGGTTACATCGGCAGCCACACCGTCTGAGCTGCTCA 230
QY 114 AGCTGGCTTACGCTTTCAATATCGACAATTTGCAATCTCCGTCATGGAAGCAGTGA 173
DB 231 GCAGGGCTTCCGCGTCTGCTGTCGACAACTCGACAACGCTCCGAGGCGCCCTCGC 290
QY 174 CCGGTGCGCCACGAGTGTGCGCTCTGCTTCTCAGAACTCCCAATTCACCCAGGGCGA 233
DB 291 CCGGTGCGCCAGCTGCGCGGCGACAGCGGC---CAACTGCTCTTCCACAGTTGA 347
QY 234 TCTCCGGAATAGGATGACTTGGAGAACTCTTCTTCCAAACCAACATTTGATGCGGTGAT 293
DB 348 CTTTCGCGACAGGACGCGTGTGTTGGACATCTTCTCGTCGACAGGTTGAGGCTGTCAT 407
QY 294 CCAGTTCTGCGCTTGAAGCGGTTGCTGAAAGGTTTGGAAAGCCCGCTGCTATTTTGA 353
DB 408 TCATTTGCTGGCTCAAGCTGTTGGGAGAGCGTGCACAGCCCTCTACTTTACTAGA 467
QY 354 TTTTAAATTTGGTTGGCAACCACTCTACGAGTTTATGCAAGTATATTAATTCAGAAA 413
DB 468 CAACAACCTGCTGGCCACCATCACCTCTCTGGAGGTGATGCTGCGAAGCGCTGCAAGAA 527

414 GATGGTTTCTCATCTGCAACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGGA 473
Db
528 GCTGGTGTCTGCTATCTGCACTGTCTATGGTGGCCCAAGGAGTACCTTGACCGGA 587
Qy 474 GGATTTCAAGTTACAGCTATGAATCCCTATGGAAGGACCAAGCTTTTCCTGGGAAGAAAT 533
Db 588 AGAATTTCCGCGCTCTGCGCCACCAATCCCTATGCGCGGACAAAGCTTGTGATTTGAAGACAT 647
Qy 534 TGCCCGAGATATTTCAGAAAGCTGACCAAGATGGAAGATCATATTACTGAGATCTTCAA 593
Db 648 CTGCCCGAGCTGCCACCGCTCGACCCCGACTGGAGATCATATCTGCTCAGTACTTCA 707
Qy 594 TCCAGTTGGGGTCTATGAAGTGGCAACTCGGTGAAGATCCCAAGGGCATCCCCAAATAA 653
Db 708 CCGCTGGCGTCTATCCAAAGGGGTACATCGCGGAAGACCCCTGCGGTGTCCCGACAA 767
Qy 654 CCTCATGCTTACATTCAGCAAGTACTGTTGGAAGATGACTGAACTCAATGTATACGG 713
Db 768 CTTGATGCTTACGTCAGCAAGTCTGTTGGAGGTTTACCTCACCTCAGCTTACGG 827
Qy 714 TCATGATTATCCAAAGGATGGCTCTGCGATCCGGGACTATATCCATGTGATGACTT 773
Db 828 GACGGACTACAGCACCAAGATGGGACTGGGTGGGTGATTTATATCCACGTTGCCGACT 887
Qy 774 GCGAGATGCCATATGCTGCCCTGCGAAAGCTCTTCA---ACGGAGACATAGTTG 830
Db 888 GGTGACGGCCACATAGAGCCCTGAGGAAGCTCTACGAAGACTCCGACAAATAGGCTG 947
Qy 831 TACTGTCTTCAACCTGGGAACCTGGTCTGGAACATCTGTGCTTGAATGGTTTACAGCAT 890
Db 948 TGAAGTGTACAACTTGGGACTGGAAGGGGAGCGTCCGTGTTGGAAATGGTGGCTGATT 1007
Qy 891 TGAAGAGCTTCTGGCAAGAAATCCAGTAAATATGTCTCAAGAGACCGGGAGATGC 950
Db 1008 CGAAGAGGCTTCTGGGAAGAAATCCCTCTGTGTGCTGGCGCAAGACCGCGAGACGC 1067
Qy 951 GACTGAGGTTTATGCACTACAGAGAGCTGAGAAAGAACTGTTGTAAGGCAAACTA 1010
Db 1068 AGAGATCGTCTACCGCGCACTGCCAAGCAGAGAGAGCTCAATGGAAGGCCAAGTA 1127
Qy 1011 TGGTGTGGAGGAGATGTGAGGACCAATGGAATTTGGGCAAGAACATCCCTGGGGTTA 1070
Db 1128 CGGGATCGAGGAGATGTGAGAGATCTGTGGAATCTGGGAGCAAGACCCGTTACGGTA 1187
Qy 1071 CGCGGGG 1077
Db 1188 CGCTGGG 1194

RESULT 13

US-10-425-114-1181
; Sequence 1181, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1181
; LENGTH: 1531
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700103191_FLI
US-10-425-114-1181

Query Match 32.8%; Score 430.2; DB 12; Length 1531;
Best Local Similarity 65.0%; Pred. No. 1.4e-106; Indels 6; Gaps 2;
Matches 668; Conservative 0; Mismatches 353;

Qy 54 CATCTGGTTCACCGTGGTCCCGTTTCATTTGGCAACCCACACCGTGTGAGTCTTCAA 113
Db 177 CATCTCGTACGCGGGCGCGGGTACATCGGCAGCCACACCGTGTGAGTCTGCA 236
Qy 114 AGCTGGTTCAGAGTTTCAATATTCAGAAATTTTCGATACTCGTTCATGGAACAGTGA 173
Db 237 GCAGGGTTCGCGGTCTGCTGCTGCAACCTCGAACCTCCAAATTCACCCAGGGCGA 236
Qy 174 CCGCGTCCGCAAGTGGTTCGCTCTGCTGCTGCAACCTCGAACCTCCAAATTCACCCAGGGCGA 233
Db 297 CCGCGTCCGCGAGCTCGCGGGCAGCAGGCGCG---CAACCTGCTCTTCCACAAGGTGA 353
Qy 234 TCTCCGAATAGGATGACTTTGGAGAACTCTTCTCCAAACAACTTTGATGCGCGTAT 293
Db 354 CTTTCGCGAGCAGGACGCTGTTGGAGAGCGTGCACAGCCCTACTTTACTACGA 413
Qy 294 CCACCTTTGCTGGCTTGAAGCGTGTGTAAGCGTGTGGAAGCGTGGAAAGCGCGTCTGATTTGA 353
Db 414 TCACCTTTGCTGGCTCAGGCTGTGGGAGAGCGTGCACAGCCCTACTTTACTACGA 473
Qy 354 TTTTAATTTGGTGGCAACATCAACCTCTAGAGTTTATGGCAAGTATATTTGCAAAA 413
Db 474 CAACAACCTGGTGGCAGCACCCTCCCTCTGGAGGTGATGGCTGCGAAAGCGTGCAGAA 533
Qy 414 GATGGTTTCTCATCTCATCTGCAACCGTTCATGCGCAACCTGAAAGCGTGGAAAGATACCGTGTGAGGA 473
Db 534 GCTGGTGTCTCTGCTCATCTGCAACTGTCTATGGTGGCCCAAGAGTACCTGTCAGGTACTCAA 593
Qy 474 GGATTTCAAGTTACAGCTATGAAATCCCTATGGAAGCGTCCCAAGGCGATCCCAATAA 533
Db 594 AGAATTTCCGCTCTGGCGCCACCAATCCCTATGCGCGGACAAAGCTTTGTGATTTGAAGACAT 653
Qy 534 TGCCCGAGATATTTCAGAAAGCTGAAACCGTTCATGCGCAACCTGAAAGCGTGGAAAGATACCGTGTGAGGA 593
Db 654 CTGGCGGAGCTGCAACCGCTCCGACCCGACTGGAGATCATACTGCTCAGGTACTCAA 713
Qy 594 TCCAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAGATCCCAAGGCGATCCCAATAA 653
Db 714 CCGCTTTGGGCTCATCCAAAGCGGTACATCGCGGAAGACCCCTGCGGTGTCCCGAACAA 773
Qy 654 CCTCATGCTTACATTCAGCAAGTGTGTTGGAAGATGACTGAACTCAATGTATACGG 713
Db 774 CTTGATGCCCTACGTGCAGCAAGTGTGTTGGAAGTACCTCACTCAGCGTCTACGG 833
Qy 714 TCATGATTATCCAAAGGAGTGGCTCTGCGATCCGGGACTATATCCATGTGATGAGCTT 773
Db 834 GACGGACTACAGCAACCAAGGATGGGACTGGGGTGGGTGATTTACATCCACGTTGTCGACCT 893
Qy 774 GCGAGATGGCCATATTGCTGCCCTGCGAAAGCTCTTCA---ACGGAGAACATAGGTTG 830
Db 894 GGTGACGGCCCATAGAGCCCTGAGGAAGCTCTAGGAAGACTTCGAAAGACTCCGACAAATAGGCTG 953
Qy 831 TACTGCTTACAACTGGGAACCTGGTCTGTTGGAACATCTGTGCTTTGAAATGGTTACAGATT 890
Db 954 TGAAGTGTACAACTGGGAGCTGGAAGGGGAGCTCCGTGTTTGGAAATGGTGGCTGATT 1013
Qy 891 TGAAGGCTTCTGGCAAGAAATTCAGTAAATATGTCCAAAGAGACCGGGAGATGC 950
Db 1014 CGAAGGCTTCTGGGAAGAAATCCCTCTGGTGTGCTGGGCGAAGACCCCGAGACGC 1073
Qy 951 GACTGAGGTTTATGCACTTACAGAGAGCTGGAAGAACTGTTGGTGGAGGCAAACTA 1010
Db 1074 AGAGATGCTCTACGCCCACTCCCAAGGCAGAGAGAGCTCAATGGAAGGCCAAGTA 1133
Qy 1011 TGGTGTGAGGAGATGTGAGGACCAATGGAATTTGGGCAAGAACTCCCTGGGGTTA 1070
Db 1134 CGGATCGAGGAGATGTGAGAGATCTGTGGAATCTGGCGAGCAAGACCCGTTACGGGTA 1193

QY 1071 CGCGGG 1077
Db 1194 CGCTGG 1200

RESULT 14
US-10-425-114-22869
; Sequence 22869, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22869
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB592-035-C3_FLI
US-10-425-114-22869

Query Match 30.8%; Score 403.8; DB 12; Length 1200;
Best Local Similarity 64.8%; Pred. No. 1.9e-99;
Matches 616; Conservative 0; Mismatches 332; Indels 3; Gaps 1;

QY 54 CATTCTGGTCAACGGGTGGTCCGGTTTCATTGGCACCCACACCGTGGTTTCAGCTTCTCAA 113
Db 253 CGTGTGGTCAACGGGTGGTCCGGTTTCATTGGCACCCACACCGTGGTTTCAGCTTCTCAA 312
QY 114 AGCTGCTTCAGGTTTCATATACCAATTCGATACCTCGTCATCGTAGCCACGGCTGCTCTCT 173
Db 313 CGCGGGCTTCCGGCGCGTGTGATCAGCACTTCAACAACTCTCGAGCTGGCGGTCCG 372
QY 174 CCCTGTCGCGCAAGTGGTGGCCCTCTGCTTTCTCAGAACCTTCAATTCACCCAGGGGGA 233
Db 373 CGCGTGGC---GCGCTCGGGGACCACTCCCGCACTCTCTTCCACAAGATTGA 429
QY 234 TCTCCGGAATAGGATGACTTGGAGAACTTCTCCTCAAAACAACATTTGATGCGGTGAT 293
Db 430 TCTCCGTGACAGGGGACACTGGAAATGGTTTGTCTTCTACAAGATTTTGAAGCTGTCTAT 489
QY 294 CCACTTTGCTGGCTTGAAGCGGTTCTGAAAGCGTTGCGAAGCCCGCTCGCTATTTTGA 353
Db 490 TCACCTTCGTGGATTGAAGCTGTGGTGAAGCGTACAGAGCCATTACTTTATTATGA 549
QY 354 TTTTAAATTTGGTGGCACCATCACTCTACAGATTATGGCAAGTATAATTTGCAAAA 413
Db 550 CAACAACGTCATTGGCACGATTAATCTCTAGAAGTTATGTCTGTTCACGGTTGCAAGAA 609
QY 414 GATGGTTTCTCATATCTGCAACCTTTATGGCCAACTGAAAGATACCGTGTGAGGA 473
Db 610 GTTGGTCTCTCATATCAGTGCAGTTTATGATCACCCTCAAACTCACCTTCACGACGA 669
QY 474 GGAATTTCAAGTTACAGCTATGAATCCCTATGACGGAACAAGCTTTTCTGGAAGAAT 533
Db 670 AAATTTTCTTACTTCAACCAATCCATATGGCAAAACAAAGCTGGTTGTTGAAGATAT 729
QY 534 TGCCCGAGATATTCAGAAAGCTGAACCAAGATGGAAGATCATATTACTGAGATATCTCAA 593
Db 730 TTCCCGGATATCTACCGTTTCACTCTGAAATGGAAGATCATTTTACTTAGTACTTCAA 789
QY 594 TCCAGTTGGGGCTTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGGATCCCAATAA 653
Db 790 TCCAGTTGGTCTCATCTTAGTGAATCTTGGCGAGGACCCACGAGGAATTTCCCAACA 849

QY 654 CCTCATGCTTACATTCAAGTAGTCTGTTGGAAGATTGACTGAACCTCAATGTATACGG 713
Db 850 TCTTATGCCCTATGTTTCAAGAGTTGGGTTGGTAGAGGCCAGCTTAACAGTTTATAG 909
QY 714 TCATGATTATCCCAAGAGGATGGCTCTGCGATCCGGGACTATATCCATGTGATGACTT 773
Db 910 AATGACTATGCAACAAGAGATGGGACTGGGTCGGGATTTACATCCATGTGTTGACCT 969
QY 774 GGCAGATGGCCATATTGCTGCCCTGCGAAAGCTCTTCCAAACGAGAACATAGTTGTAC 833
Db 970 TGCTGACGGACATATTGCTGCTTTCGAGAAGCTTTTGAGAATCTCTAGCATAGGTTGA 1029
QY 834 TGCTTACAACTGGGAACTGGTGGTGGAAATCTGTGCTTGTGAAATGTTTACAGCATTTGA 893
Db 1030 AGCTGATCAACCTTGGAAACCGGAAAGGATCATCTGTCTGCGAGATTGTTAAAGCATTTGA 1089
QY 894 AAGGCTTCTGGCAAGAAATTCAGTAAATATATGTCCAAGAACCGGGAGATGCGAC 953
Db 1090 GAGGCTTCTGGGAAGAAATATCTCTGATTTTGGTGAAGACGCCAGGTGATGCGAGA 1149
QY 954 TGAGTTTATGCTATCTACAGAGAGCTGAGAAGAACTTTGGTTGGAAGC 1004
Db 1150 GATTCTGTTTTCAGAGACTACTAAAGCAGAGAGGAGCTTAACCTGGAAGC 1200

RESULT 15
US-10-425-114-3333
; Sequence 3333, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3333
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700260760_FLI
US-10-425-114-3333

Query Match 30.7%; Score 402.2; DB 12; Length 1729;
Best Local Similarity 65.3%; Pred. No. 6.7e-99;
Matches 590; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 170 TGGACCGCTCCGCAAGTAGTGGTGGCCCTCTGCTTTCTGAGAAGCTTCAATTCACCCAGG 229
Db 133 TGCAGCTCTCTCTCGCGGGCTTCGCGCCGCTCGTCATCGACAACCTCTCTTCCACAAGA 192
QY 230 GCGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAAACAACATTTGATGCGG 289
Db 193 TTGATCTCCGTCAGACAGGAGACTGGAATGGTTTGTCTTCTACAAGATTTGAAGCTG 252
QY 290 TGATCCATTTGCTGGCTTGAAGCGGTTGCTGAAAGCGTTGGAAGCGTGGGACCGCGTCTATT 349
Db 253 TCATTCTCTCGCTGGATTGAAAGCTGTGGGTGAAAGCTGACAGAACCTTACTTTTATT 312
QY 350 TTGATTTTAAATTTGGTGGCACCATCAACCTCTACAGATTTATGGCAAGATATAATTGCA 409
Db 313 ATGACAACAGCTATTGGCACGATTAATCTTCTAGAGTTATGTCTGTTCACGGTTGCA 372
QY 410 AAAAGATGGTTTCTCATCATCTGCAACCGTTTATGSCCAACCTGAAAAGATACCGGTGTG 469

```

Db      373 AGAAGTGTGTCTCTCATCATCAGCTGCAGTGTATGATCAACCCAAAAAACAACCCCTGCA 432
QY      470 AGGAGGATTTCAAGTTTACAAAGTATGAATCCCTATGAGACGGACCAAGCTTTTCTGTGAAG 529
Db      433 CAGAAATTTTCTCTTCTTCTCABACATCCATATGGCAAAACAAGCTCGTTGTGAAG 492
QY      530 AAATGCCCGAGATATTAGAAAGCTGAACAGAGATGGAGATCATATATCTAGATACT 589
Db      493 ATATTTGCCGGGATATCTACCGTTCCAGATCCCTGAATGGAGATCATTTTACTTAGGTACT 552
QY      590 TCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAA 649
Db      553 TCATCCAGTTGGTGTCTCATCTAGTGGATATCTTGGCGAGGACCCACGAGGATTTCCA 612
QY      650 ATAACTCATGCTTACATTCAGCAAGTAGCTGTGGAGATTTGAACTCAATGTAT 709
Db      613 ACAATCTTATGCCCTATGTTTCAAGAGTTGCGGTTGGTAGAGGCCAGCTCTAACAGTTT 672
QY      710 ACGGTCTATGATTTATCAACGAGGATGGCTCTGCGATCCGGGACTATATCCATGTGATGG 769
Db      673 TAGGAAATGACTATGCAACAGAGATGGGACTGGGTCGAGATTTACATCCATGTGGTTG 732
QY      770 ACTTGGCAGATGGCCATATTGTGCCCTCGAAAGCTCTTCAACGGAGAACATAGGTT 829
Db      733 ACCTTGCTGACGGACATATTGCTGCATTGCAAGGCTTTTGAGAACTCTAGCATAGGGT 792
QY      830 GTACTGCTTACACCTGGGAACTGGTCTGTGGACATCTGTGCTTGAATGGTTACAGCAT 889
Db      793 GTGAGCGTACACCTTGGAAACCGAAGAGGTACATCTGTCTGGAGATTGTTAAAGCAT 852
QY      890 TTGAAAGGCTTCTGGCAAGAAATTTCCAGTAAATTTATGTCCAAGAGACCGGGAGATG 949
Db      853 TTGAGAAGGCTTCTGGGAAGAAATACCTCTGATTTTGGTGAAGAAGCGCCAGGTGATG 912
QY      950 CGACTGAGGTTTATGCTATCTACAGAGAGCTGAGAAAGACTTGGTTGGAGGCAAACT 1009
Db      913 CAGAGATTCTGTTTTCAGAGACTACTAAAGCAGAGAGGGAGCTTAACCTGGAAGCAAAAT 972
QY      1010 ATCGTGTGGAGGAGATGTGAGGAGCAATGGAATTTGGCAAGAACCAATCCCTGGGGTT 1069
Db      973 ACGGTATTGAGAGATGTGCGCGACCAATGGAACTGGGCGCAGCAAGAACCTTATGGCT 1032
QY      1070 ACG 1072
Db      1033 ATG 1035

```

Search completed: March 22, 2004, 21:52:30
 Job time : 498 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 17:30:18 ; Search time 1766 Seconds

(without alignments)
10403.394 Million cell updates/sec

Title: US-09-913-064A-13

Perfect score: 1312

Sequence: 1 gacagagccactctctctcc.....aaaaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:**

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_estl:**

9: gb_est1:**

10: gb_est2:**

11: gb_est3:**

12: gb_est4:**

13: gb_est5:**

14: gb_est6:**

15: em_estfun:**

16: em_eston:**

17: em_gss_hum:**

18: em_gss_inv:**

19: em_gss_pln:**

20: em_gss_vrt:**

21: em_gss_fun:**

22: em_gss_mam:**

23: em_gss_mus:**

24: em_gss_pro:**

25: em_gss_rod:**

26: em_gss_phg:**

27: em_gss_vrl:**

28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	765.4	58.3	775	14	CA802698 sau40g01.
2	636.8	48.5	664	10	AW508396 si40d07.y
3	599.2	45.7	649	10	BF324112 su22a04.y
4	584.8	44.6	613	10	BF324111 su22g12.y

5	573.4	43.7	575	12	BM521748
6	571.4	43.5	574	12	BM086182
7	562.8	42.9	567	10	AW472274
8	561.4	42.8	564	10	AW569126
9	560.2	42.7	798	14	CB892692
10	558.8	42.6	580	10	AW507982
11	558.4	42.6	564	9	AI856802
12	550.8	42.0	560	10	BE474654
13	540.6	41.2	547	12	BM528276
14	535.2	40.8	694	10	BF635019
15	532	40.5	552	12	BM891597
16	531	40.5	718	12	EG644768
17	530.4	40.4	879	13	BQ797360
18	525.6	40.1	718	12	EG647665
19	519.6	39.6	685	12	EG455279
20	518.4	39.5	564	12	BI498305
21	510	38.9	556	10	BE352750
22	509.6	38.8	551	12	BI946267
23	502	38.3	690	10	AW775957
24	501	38.2	613	12	BM813737
25	498.2	38.0	648	12	BI271421
26	495.2	37.7	626	12	BI270061
27	491.4	37.5	493	12	BI944852
28	491.2	37.4	560	12	BM177819
29	490	37.3	526	12	BI942090
30	489.6	37.3	721	14	CF209053
31	484.4	36.9	812	12	EG584474
32	481.4	36.7	651	10	AW774538
33	479.2	36.5	509	10	BE661026
34	478.2	36.4	753	13	BU635732
35	468.6	35.7	579	14	CB894137
36	465.2	35.5	672	10	BF645662
37	454.2	34.6	583	10	AW507725
38	451.8	34.4	475	10	AW507725
39	450	34.3	566	12	BI270893
40	447.2	34.1	719	14	CD841689
41	443	33.8	1816	11	AY104881
42	434	33.1	899	14	CK284851
43	433.4	33.0	727	13	BU018048
44	431	32.9	431	12	BI943653
45	429.6	32.7	437	10	AW397514

ALIGNMENTS

RESULT 1
CA802698 775 bp mRNA linear EST 05-DEC-2002
LOCUS sau40g01.y1 Gm-cl071 Glycine max cDNA clone SOYBEAN CLONE ID:
DEFINITION Gm-cl071-3386 5', similar to TR:Q43070 Q43070
UDP-GALACTOSE-4-EPIMERASE ; mRNA sequence.

CA802698 GI:26059784

VERSION EST.

KEYWORDS

SOURCE Glycine max (soybean)

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 775)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,

Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,

Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCaum, R., Waterston, R., and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

Public Soybean EST Project

Washington University School of Medicine

CA802698 775 bp mRNA linear EST 05-DEC-2002
sau40g01.y1 Gm-cl071 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl071-3386 5', similar to TR:Q43070 Q43070
UDP-GALACTOSE-4-EPIMERASE ; mRNA sequence.

CA802698

GI:26059784

EST.

Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 775)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,

Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,

Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCaum, R., Waterston, R., and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: ccu@resgen.com web site:
 www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 456.
 Location/Qualifiers

FEATURES

1..775
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl071-3386"
 /tissue type="Immature pods (~2cm long) of greenhouse
 grown plants"
 /lab_host="DH10B"
 /clone_lib="Gm-cl071"
 /notes="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; The
 cDNA library was constructed from mRNA isolated from
 immature pods (approximately 2cm long) of greenhouse grown
 plants. The library was prepared using the Life
 Technologies pSuperScript cDNA library construction kit.
 Complementary DNA was synthesized from mRNA using a
 poly(dT) sequence with a NotI restriction site. SalI
 linkers adapters were ligated to the blunt-ended cDNA
 fragments followed by NotI digestion. The cDNA fragments
 were directionally cloned into the NotI-SalI restriction
 site of the pSPORT1 vector. The ligated cDNA fragments
 were transformed into E.coli Electromax DH10B host cells.
 This library was constructed in the laboratory of Dr. Lila
 Vodkin by Anu Khanna at the University of Illinois at
 Urbana-Champaign. email: l-vodkin@uiuc.edu"

ORIGIN

Query Match 58.3%; Score 765.4; DB 14; Length 775;
 Best Local Similarity 99.2%; Pred. No. 2.8e-121;
 Matches 769; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 519 TTTCCTGGAGAAATTGCCCGAGATATTTCAGAAAGCTGACACGAGTGAAGATCATATT 578
 Db 481 TTTCCTGGAGAAATTGCCCGAGATATTTCAGAAAGCTGACACGAGTGAAGATCATATT 540
 Qy 579 ACTGAGATACCTTCAATCCAGTTCGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAA 638
 Db 541 ACTGAGATACCTTCAATCCAGTTCGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAA 600
 Qy 639 GGGCATCCCAATAAAGCTGATGCTTACATTCAGCAAGTAGCTGTTGGAGATTGACTGA 698
 Db 601 GGGCATCCCAATAAAGCTGATGCTTACATTCAGCAAGTAGCTGTTGGAGATTGACTGA 660
 Qy 699 ACTCAATGTATACGGTTCATGATTATCCACAGGGATGGCTCTGCGATCCGGGACTATAT 758
 Db 661 ACTCCATGTATACGGTTCATGATTATCCACAGGGATGGCTCTGCGATCCGGGACTATAT 720
 Qy 759 CCATGTGATGAGCTTGCGAGATGCGCAATATGCTGCTCGCAAGCTCTTCACA 813
 Db 721 CCATGTGATGAGCTTGCGAGATGCGCAATATGCTGCTCGCAAGCTCTTCACA 775

AW508396 664 bp mRNA linear EST 03-DEC-2001
 s140607.y1 Gm-r1030 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-r1030-1382 5' similar to TR:Q43070 Q43070
 UTP-GALACTOSE-4-EPIMERASE ; mRNA sequence.

AW508396 GI:7146474
 AW508396.1
 Glycine max (soybean)
 EST.
 Glycine max
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 1 (bases 1 to 664)
 Shoemaker,R., Keim,P., Vodkin,L., Rpeiding,J., Coryell,V.,
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,I., Martin,J.,
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
 Schurr,K., Ritter,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 Insert Length: 1309 Std Error: 0.00
 High quality sequence stop: 441.
 Location/Qualifiers
 1..664
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-r1030-1382"
 /lab_host="DH10B"
 /clone_lib="Gm-r1030"
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
 cDNA library was constructed from mRNA isolated from
 immature cotyledons of greenhouse grown plants
 (individual seed fresh weight of 100-300mg). The library
 was prepared using the Life Technologies pSuperScript cDNA
 library construction kit. Complementary DNA was
 synthesized from mRNA using a poly(dT) sequence with a
 NotI restriction site. SalI linkers adapters were ligated

TITLE
 JOURNAL
 COMMENT

FEATURES
 source

to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSP0rt1 vector. The ligated cDNA fragments were transformed into E. coli ElectroMax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note that Gm-r1030 is a re-rack of Gm-cl007."

ORIGIN	
Query Match	48.5%; Score 636.8; DB 10; Length 664;
Best Local Similarity	97.4%; Pred. No. 3e-99;
Matches	647; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy	30 CAGATGGTGTCTTCCTCCCAACACATCTGGTCACCGGTGGTGGCGTTTCATTTGGCAC 89
Db	1 CAGATGGTGTCTTCCTCCCAACACATCTGGTCACCGGTGGTGGCGTTTCATTTGGCAC 60
Qy	90 CCACACCGTCTTCAGCTTCTCAAGCTTCCAGCTTCCAGCTTTCATATCGACAAATTCGA 149
Db	61 CCACACCGTCTTCAGCTTCTCAAGCTTCCAGCTTCCAGCTTTCATATCGACAAATTCGA 120
Qy	150 TAATCTCGTCAATGAAGCAGTGGACCGGTCCGCAAGTGGTGGCGCTCTGCTTTCTCA 209
Db	121 TAATCTCGTCAATGAAGCAGTGGACCGGTCCGCAAGTGGTGGCGCTCTGCTTTCTCA 180
Qy	210 GAACCTCCAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAAATCTTCTC 269
Db	181 GAACCTCCAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAAATCTTCTC 240
Qy	270 CAAAACAACTTTGATGCCGTGATCCACTTTTGTGGCTTGAAGCGTGTCTGCAAGCGT 329
Db	241 CAAAACAACTTTGATGCCGTGATCCACTTTTGTGGCTTGAAGCGTGTCTGCAAGCGT 300
Qy	330 TGGAGAGCCCGTCTGCTATTTTGAATTTTAAATTTGGTTGGCACCATCAACCTCTACGATT 389
Db	301 TGGAGAGCCCGTCTGCTATTTTGAATTTTAAATTTGGTTGGCACCATCAACCTCTACGATT 360
Qy	390 TATGCGCAAGTATATATGCAAAAGATGGTTTCTCATCTATCTGCAACCGTTTATGCGCA 449
Db	361 TATGCGCAAGTATATATGCAAAAGATGGTTTCTCATCTATCTGCAACCGTTTATGCGCA 420
Qy	450 ACCTGAAAGATACCGTGTGAGAGGATTTCAAGTTCAAGTATGATCCCTATGGAAG 509
Db	421 ACCTGAAAGATACCGTGTGAGAGGATTTCAAGTTCAAGTATGATCCCTATGGAAG 480
Qy	510 GACCAAGCTTTCTGGAAGAAATTCGCCGAGATATTCAGAAAGCTCAACCAAGTGGAA 569
Db	481 GACCAAGCTTTCTGGAAGAAATTCGCCGAGATATTCAGAAAGCTCAACCAAGTGGAA 540
Qy	570 GATCATATTACTGAGTACTTCAATCCAGTTGGGGCTCATGAAAGTGGCAACTCGGTGA 629
Db	541 GATCATATTACTGAGTACTTCAATCCAGTTGGGGCTCATGAAAGTGGCAACTATGTGA 600
Qy	630 AGATCCCAAGGGCATCCCAATACCTCATGCTTACATTCAGCAAGTAGCTGTGGAG 689
Db	601 AGATCCCAAGGGCATCCCAATACCTTATGCTTACATTCAGCAAGTAGCTGTGGAG 660
Qy	690 ATTG 693
Db	661 AATG 664

RESULT 3
BF324112
LOCUS
DEFINITION
su22h04.v1 Gm-cl068 Glycine max cDNA clone GENEOME SYSTEMS CLONE ID:
Gm-cl068-343 5', similar to TR:Q43070 Q43070
UDP-GALACTOSE-4-EPIMERASE ;, mRNA sequence.
ACCESSION
BF324112
VERSION
BF324112.1 GI:11273736
KEYWORDS
EST.
SOURCE
Glycine max (soybean)
ORGANISM
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 649)
Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 420.

Location/Qualifiers
1. 649
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl068-343"
/tissue_type="Leaf, drought stressed, 1 month old plants, greenhouse grown"
/lab_host="DH10B"
/clone_lib="Gm-cl068"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from drought stressed leaf tissue of the cultivar Williams 82. The month old greenhouse grown plants were deprived of water for 3 days prior to harvesting the stressed leaf tissue. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN	
Query Match	45.7%; Score 599.2; DB 10; Length 649;
Best Local Similarity	98.3%; Pred. No. 8e-93;
Matches	637; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
Qy	16 CTCCTCTCTATTGACGATGGTGTCTTCTCCCAACACATTCGGTCACCGGTGGTGGC 75
Db	2 CTCCTCTCTATTGACGATGGTGTCTTCTCCCAACACATTCGGTCACCGGTGGTGGC 61
Qy	76 GGTTCATTGGCACCACACCGCTCGTTTCAGTTTCCTCAAGCTGGCTTCAGCGTTTCAATA 135
Db	62 GGTTCATTGGCACCACACCGCTCGTTTCAGTTTCCTCAAGCTGGCTTCAGCGTTTCAATA 121
Qy	136 ATCGCAATTTTCGTAATCTCGTCTATGAGCAGTGGACCGCTCCGCAAGTGGTGGC 195
Db	122 ATCGCAATTTTCGTAATCTCGTCTATGAGCAGTGGACCGCTCCGCAAGTGGTGGC 181
Qy	196 CCTCTGCTTTCTCAGAACCTTCAATTCACCCAGGCGCATCTCCGAAATAGGATGACATTG 255
Db	182 CCTCTGCTTTCTCAGAACCTTCAATTCACCCAGGCGCATCTCCGAAATAGGATGACATTG 241
Qy	256 GAGAAACTCTTCTCCAAACACACATTTGATCCCGTGATCCACTTTGCTGGCTTGAAGCG 315
Db	242 GAGAAACTCTTCTCCAAACACACATTTGATCCCGTGATCCACTTTGCTGGCTTGAAGCG 301

QY 316 GTTCTGGAAGCGTTCGGAAGCCCGTCGCTATTTTGAATTTTAAATTTGGTTGGCACCATC 375
 |||||
 Db 302 GTTCTGGAAGCGTTCGGAAGCCCGTCGCTATTTTGAATTTTAAATTTGGTTGGCACCATC 361
 |||||
 QY 376 AACCTTACGAGTTTATGCAAAAGTATAATTCGCAAAAGATGGTTTCTCATCATCTGCA 435
 |||||
 Db 362 AACCTTACGAGTTTATGCAAAAGTATAATTCGCAAAAGATGGTTTCTCATCATCTGCA 421
 |||||
 QY 436 ACCGTTTATGGCAACCTCGAAGATACCGTGTGAGGAGATTCAAGTTTACAAGCTATG 495
 |||||
 Db 422 ACCGTTTATGGCAACCTCGAAGATACCGTGTGAGGAGATTCAAGTTTACAAGCTATG 481
 |||||
 QY 496 AATCCCTATGGACGACCAAGCTTTTCTCGAAGAAATTCGCCGAGATATTCAGAAAGCT 555
 |||||
 Db 482 AATCCCTATGGACGACCAAGCTTTTCTCGAAGAAATTCGCCGAGATATTCAGAAAGCT 541
 |||||
 QY 556 GAA-CCAGATGGAAGATCATATTAC-TGAGTACTTCAATCCAG-TTGGGGCTCATGAA 612
 |||||
 Db 542 GAAACCAGATGGAAGATCATATTACATGAAATTTCTCAATACAGTTTGGGGCTCATGAA 601
 |||||
 QY 613 AGTGCGAAATCTCGTGAAGATCCCAAGGGCATCCCAAAATTAACCTCATG 660
 |||||
 Db 602 AGTGCGAAATCTCGTGAAGATCCCAAGGGCATCCCAAAATTAACCTCATG 649
 |||||

RESULT 4
 BF324111 613 bp mRNA linear EST 06-DEC-2001
 LOCUS su22g12.v1 Gm-c1068 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 DEFINITION UDP-GALACTOSE-4-EPIMERASE ;, mRNA sequence.

ACCESSION BF324111
 VERSION BF324111.1 GI:11273735
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 613)
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Shin, I., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 CONTACT: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 5501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

COMMENTS This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 High quality sequence stop: 365.
 Location/Qualifiers
 1..613
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1068-311"
 /tissue_type="Leaf, drought stressed, 1 month old plants,
 greenhouse grown"
 /lab_host="DH10B"
 /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; The cDNA library was constructed from mRNA isolated

FEATURES

source
 RESULT 5
 BM521748
 LOCUS
 DEFINITION sak70f12.v1 Gm-c1036 Glycine max cDNA clone linear
 EST 15-FEB-2002
 BM521748 575 bp mRNA
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1036-311"
 /tissue_type="Leaf, drought stressed, 1 month old plants,
 greenhouse grown"
 /lab_host="DH10B"
 /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; The cDNA library was constructed from mRNA isolated

from drought stressed leaf tissue of the cultivar Williams
 82. The month old greenhouse grown plants were deprived of
 water for 3 days prior to harvesting the stressed leaf
 tissue. Complementary DNA was synthesized from mRNA using
 a primer consisting of a poly(dT) sequence with a XhoI
 restriction site. EcoRI adapters were ligated to the
 blunt-ended cDNA fragments followed by XhoI digestion. The
 cDNA fragments were directionally cloned into the
 EcoRI-XhoI restriction site of the pBluescript vector. The
 ligated cDNA fragments were transformed into DH10B host
 cells (GibcoBRL). This library was constructed in the
 laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 44.6%; Score 584.8; DB 10; Length 613;
 Best Local Similarity 97.3%; Pred. No. 2.4e-90;
 Matches 595; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 QY 24 CTATTGAGCATGGTGTCTTCTCCCAACATCTTGGTCACCGGTGGTCCGGTTTCAT 83
 Db 1 CTATTGAGCATGGTGTCTTCTCCCAACATCTTGGTCACCGGTGGTCCGGTTTCAT 60
 |||||
 QY 84 TGGCACCACACACCGTCGTTTCAGCTTCTCAAAGCTGGCTTCAGCGTTTCAATAATCGACAA 143
 Db 61 TGGCACCACACACCGTCGTTTCAGCTTCTCAAAGCTGGCTTCAGCGTTTCAATAATCGACAA 120
 |||||
 QY 144 TTTCGATPAATCCCGTCATGAAGCAGTGGACCGCGTCCGCCAAGTGGTTGGCCCTCTGCT 203
 Db 121 TTTCGATPAATCCCGTCATGAAGCAGTGGACCGCGTCCGCCAAGTGGTTGGCCCTCTGCT 180
 |||||
 QY 204 TTCTCAGAACTCTCAATTCACCCAGCGGCGATCTCCGGAATAGGATGACTTTGGAGAACT 263
 Db 181 TTCTCAGAACTCTCAATTCACCCAGCGGCGATCTCCGGAATAGGATGACTTTGGAGAACT 240
 |||||
 QY 264 CTTCTCCAAAACAACTTTGATGCCGTGATPCCAATTTGGCTGGCTTGAAGCGTTGCTGA 323
 Db 241 CTTCTCCAAAACAACTTTGATGCCGTGATPCCAATTTGGCTGGCTTGAAGCGTTGCTGA 300
 |||||
 QY 324 AAGCGTTGCCAAGCCCGTCGCTATTGATTTTAAATTTGGTTGGCACCACCACTCTA 383
 Db 301 AAGCGTTGCCAAGCCCGTCGCTATTGATTTTAAATTTGGTTGGCACCACCACTCTA 360
 |||||
 QY 384 CGAGTTTATGGCAAGATATAATTCGAAAAGATGGTTTCTCATCATCTGCAACCGTTTA 443
 Db 361 CGAATTTATGGCAAGATATAATTCGAAAAGATGGTTTCTCATCATCTGCAACCGTTTA 420
 |||||
 QY 444 TGGCCAACTGAAAGATACCGTGTGAGGAGGATTTCAAGTTACAGCTATGATCCCTA 503
 Db 421 TGGCCAACTGAAAGATACCGTGTGAGGAGGATTTCAAGTTACAGCTATGATCCCTA 480
 |||||
 QY 504 TGGACGGACCAAGCTTTCTCGAAGAAATTCGCCGAGATATTCAGAAAAGCTGAACACAG 563
 Db 481 TGGACGGACCAAGCTTTCTCGAAGAAATTCGCCGAGATATTCAGAAAAGCTGAACACAG 540
 |||||
 QY 564 ATGAGATCATATTTACTGATCTTCAATCCAGTTGGGCTCATGAAAGTGGCAACT 623
 Db 541 ATGAGATCATATTTACTGATCTTCAATCCAGTTGGGCTCATGAAAGTGGCAACT 600
 |||||
 QY 624 CGGTGAAGATCC 635
 Db 601 CGATGAATCCC 612
 |||||

RESULT 5

BM521748
 LOCUS
 DEFINITION sak70f12.v1 Gm-c1036 Glycine max cDNA clone linear
 EST 15-FEB-2002
 BM521748 575 bp mRNA
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1036-311"
 /tissue_type="Leaf, drought stressed, 1 month old plants,
 greenhouse grown"
 /lab_host="DH10B"
 /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; The cDNA library was constructed from mRNA isolated

ORGANISM Glycine max

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatoxyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine

AUTHORS 1 (bases 1 to 575)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project

JOURNAL Unpublished (1999)

COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com

Seq primer: -4ORP from Gibco

High quality sequence stop: 435.

Location/Qualifiers

1..575

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl036-8352"

/tissue_type="somatic embryos cultured on MSD 20"

/lab_host="DH10B"

/clone_lib="Gm-cl036"

/notes="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Query Match 43.7%; Score 573.4; DB 12; Length 575;
Best Local Similarity 99.8%; Pred. No. 2.2e-88;
Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 50 AACACATTCCTGGTCACCGGTGGTGGCGGTTTCATTTGGCACCCACACCGCTTCAGCTTC 109

DB 1 AACACATTCCTGGTCACCGGTGGTGGCGGTTTCATTTGGCACCCACACCGCTTCAGCTTC 60

QY 110 TCAAGCTGGCTTCACGGTTTCATTAATCGACATTCGATACTCCGTCGGAAGCAG 169

DB 61 TCAAGCTGGCTTCACGGTTTCATTAATCGACATTCGATACTCCGTCGGAAGCAG 120

QY 170 TGACACCGCTCCGCCAAGTGGTGGCGCTTCGTTCTCAGAACCTCCCAATTCACCCAGG 229

DB 121 TGACACCGCTCCGCCAAGTGGTGGCGCTTCGTTCTCAGAACCTCCCAATTCACCCAGG 180

QY 230 GCATCTCCGGAATAGGATGATTCGGAGAACTCTCTCCAAAACAAATTTGATGCCG 289

DB 181 GCGATCTCCGGAATAGGATGATTCGGAGAACTCTCTCCAAAACAAATTTGATGCCG 240

QY 290 TGATCCACTTTGCTGGCTTGAAGCGGTGCTGAAAGCGTTGGGAAGCCCGCTGCTATT 349

DB 241 TGATCCACTTTGCTGGCTTGAAGCGGTGCTGAAAGCGTTGGGAAGCCCGCTGCTATT 300

QY 350 TTGATTTTAAATTTGGTTGGCACCACCTCAACCTCTACGAGTTTATGGCAAAAGTATAATTGCA 409

DB 301 TTGATTTTAAATTTGGTTGGCACCACCTCAACCTCTACGAGTTTATGGCAAAAGTATAATTGCA 360

QY 410 AAAAGATGGTTTTCATCATCTGCAACCGTTTATGGCCAACTGAAAAGATACCGTGTG 469

DB 361 AAAAGATGGTTTTCATCATCTGCAACCGTTTATGGCCAACTGAAAAGATACCGTGTG 420

QY 470 AGGAGGATTCAGTTTACAGCTATGAATCCCTATGACCGGACCAAGCTTTCTCTGGAAG 529

DB 421 AGGAGGATTCAGTTTACAGCTATGAATCCCTATGACCGGACCAAGCTTTCTCTGGAAG 480

QY 530 AAATGGCCCGAGATATTGAGAAAGCTGAACCAAGATGAAGATCATATTACTGAGATACT 589

DB 481 AAATGGCCCGAGATATTGAGAAAGCTGAACCAAGATGAAGATCATATTACTGAGATACT 540

QY 590 TCAATCCAGTTGGGCTCATGAAAGTGGCAAACTC 624

DB 541 TCAATCCAGTTGGGCTCATGAAAGTGGCAAACTC 575

RESULT 6

LOCUS BM086182

DEFINITION BM086182 574 bp mRNA linear EST 19-NOV-2001

DESCRIPTION sahs6e02.y1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl036-3651 5' similar to TR:Q43070 Q43070

ACCESSION UDP-GALACTOSE-4-EPIMERASE ;, mRNA sequence.

VERSION BM086182

KEYWORDS EST..

SOURCE Glycine max (soybean)

ORGANISM Glycine max

REFERENCE 1 (bases 1 to 574)

AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project

JOURNAL Unpublished (1999)

COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com

High quality sequence stop: 434.

Location/Qualifiers

1..574

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl036-3651"

/tissue_type="somatic embryos cultured on MSD 20"

/lab_host="DH10B"

/clone_lib="Gm-cl036"

/note="vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20"

cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linker adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Query Match 43.6%; Score 571.4; DB 12; Length 574;
Best Local Similarity 99.7%; Pred. No. 4.8e-88;
Matches 572; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 63 CACCGTGTGGCGGTTTCATTGGACCCACACCGCTCGTTCAGCTTCTCAAGCTGGCTT 122
DB 1 CACCGTGTGGCGGTTTCATTGGACCCACACCGCTCGTTCAGCTTCTCAAGCTGGCTT 60

QY 123 CAGCGTTCAATATGCAATTTTCGATAACTTCCTCATGGAAGCAGTGGACCGGTCG 182
DB 61 CAGCGTTTCATATATGCAATTTTCGATAACTTCCTCATGGAAGCAGTGGACCGGTCG 120

QY 183 CCAAGTGTGTGGCGGTTTCATTGGACCCACACCGCTCGTTCAGCTTCTCAAGCTGGCTT 242
DB 121 CCAAGTGTGTGGCGGTTTCATTGGACCCACACCGCTCGTTCAGCTTCTCAAGCTGGCTT 180

QY 243 TAGGATGACTTTGGAGAACTCTTCTCCAAAACACATTTGATGCGGTGATCCACTTGC 302
DB 181 TAGGATGACTTTGGAGAACTCTTCTCCAAAACACATTTGATGCGGTGATCCACTTGC 240

QY 303 TGGCTTGAAGCGGTGTGGAAGCGTTCGGAAGCGGTCGCTATTTTGAATTTAATTT 362
DB 241 TGGCTTGAAGCGGTGTGGAAGCGTTCGGAAGCGGTCGCTATTTTGAATTTAATTT 300

QY 363 GGTGTGCACCATCAACCTCTACGAGTTTATGCAAGATATATTCGAAAAAGATGGTTT 422
DB 301 GGTGTGCACCATCAACCTCTACGAGTTTATGCAAGATATATTCGAAAAAGATGGTTT 360

QY 423 CTCATCATCTGAACGGTTTATGGCAACCTGAAAGATACCGTGTGAGGAGATTCAA 482
DB 361 CTCATCATCTGAACGGTTTATGGCAACCTGAAAGATACCGTGTGAGGAGATTCAA 420

QY 483 GTTACAGCTATGAATCCCTATGGACGACCAAGCTTTCTCGGAAGAAATGCCCCGAGA 542
DB 421 GTTACAGCTATGAATCCCTATGGACGACCAAGCTTTCTCGGAAGAAATGCCCCGAGA 480

QY 543 TATTGAGAAGCTGAACCAAGATGAAGATCATATTTACTGAGATACCTTCAATCCAGTTGG 602
DB 481 TATTGAGAAGCTGAACCAAGATGAAGATCATATTTACTGAGATACCTTCAATCCAGTTGG 540

QY 603 GGTCTATGAAGTGGCAACTCGGTGAAGATCCC 636
DB 541 GGTCTATGAAGTGGCAACTCGGTGAAGATCCC 574

RESULT 7

AW472274 567 bp mRNA linear EST 03-DEC-2001
LOCUS si23b06.y1 Gm-cl029 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl029-2028 5' similar to TR:Q43070 Q43070
UDP-GALACTOSE-4-EPIMERASE 1; mRNA sequence.
ACCESSION AW472274
VERSION AW472274.1 GI:7042380
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 567)
Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63109, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1331 Std Error: 0.00
High quality sequence stop: 426.

FEATURES
source

1..567
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl029-2028"
/tissue_type="very young cotyledons of greenhouse grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl029"
/notes="Vector: pSPORT1; Site 1: Not I; Site 2: Sal I; This cDNA library was constructed from mRNA isolated from very young cotyledons (20-50mgs fresh weight) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

ORIGIN

Query Match 42.9%; Score 562.8; DB 10; Length 567;
Best Local Similarity 99.8%; Pred. No. 1.4e-86;
Matches 564; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 67 GGTGGTCCCGGTTTCATTGGACCCACACCGTTCAGCTTCTCAAAAGCTGGCTTCAGC 126
DB 1 GGTGGTCCCGGTTTCATTGGACCCACACCGTTCAGCTTCTCAAAAGCTGGCTTCAGC 60

QY 127 GTTTCATATATGCACAATTTTCGATCTCCGTCATGAGACGATGGACCGCTCCGCCAA 186
DB 61 GTTTCATATATGCACAATTTTCGATCTCCGTCATGAGACGATGGACCGCTCCGCCAA 120

QY 187 GTGGTTGGCCCTCTGCTTTCTCAGAACCTCCCAATTCACCCAGGCGGATCTCCGGAATAG 246
DB 121 GTGGTTGGCCCTCTGCTTTCTCAGAACCTCCCAATTCACCCAGGCGGATCTCCGGAATAG 180

QY 247 GATGACTTGGAGAACTCTTCTCCAAAACACATTTTATGCGCGTGTATCCACCTTGTGTCG 306
DB 181 GATGACTTGGAGAACTCTTCTCCAAAACACATTTTATGCGCGTGTATCCACCTTGTGTCG 240

QY 307 TTGAAAGCGGTGCTGCTGAAGCGTTGCGAAGCCCGCTCGCTATTTTGAATTTTAAATTTGGTT 366
DB 241 TTGAAAGCGGTGCTGCTGAAGCGTTGCGAAGCCCGCTCGCTATTTTGAATTTTAAATTTGGTT 300

QY 367 GGCACCATCAACCTCTACGAGTTTATGGCAAGATATAATTCGAAAAAGATGGTTTCTCA 426

```

Db 301 GGACCATCAACCTCTACGAATTATATGGCAAGTATAATTGCAAAAGATGGTTTCTCA 360
Qy 427 TCATCTGCAACCGTTTATGGCCCAACCTGAAAGATACCGGTGTCAGAGAGATTTCAAGTTA 486
Db 361 TCATCTGCAACCGTTTATGGCCCAACCTGAAAGATACCGGTGTCAGAGAGATTTCAAGTTA 420
Qy 487 CAGCTATGATCCCTATGACGACCAAGCTTTCTGGAAGAAATTTGCCGAGATAT 546
Db 421 CAAGCTATGATCCCTATGACGACCAAGCTTTCTGGAAGAAATTTGCCGAGATAT 480
Qy 547 CAGAAAGCTGAACAGAGATGGAAGATCATATTACTGAGATACATTCATCCAGTTGGGCT 606
Db 481 CAGAAAGCTGAACAGAGATGGAAGATCATATTACTGAGATACATTCATCCAGTTGGGCT 540
Qy 607 CATGAAAGTGGCAACTCGGTGAAGA 632
Db 541 CATGAAAGTGGCAACTCGGTGAAGA 566

RESULT 8
AW569126
LOCUS
DEFINITION
S163f08.y1 Gm-r1030 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-r1030-3616 5' similar to TR:Q43070 Q43070
UDP-GALACTOSE-4-EPIMERASE ; mRNA sequence.
ACCESSION
AW569126
VERSION
AW569126.1 GI:7233783
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 564)
Shoemaker,R., Keim,P., Vodkin,L., Brpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurr,R., Ritter,B., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1359 Std Error: 0.00
High quality sequence stop: 437.
Location/Qualifiers
1..564
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-r1030-3616"
/lab_hosts="DH108"
/clone_lib="Gm-r1030"
/notes="Vector: pSPORt1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants
(individual seed fresh weight of 100-300mg). The library
was prepared using the Life Technologies pSuperScript cDNA
library construction kit. Complementary DNA was
synthesized from mRNA using a poly(dT) sequence with a
NotI restriction site. SalI linkers adapters were ligated
to the blunt-ended cDNA fragments followed by NotI

```

```

ORIGIN
Query Match 42.8%; Score 561.4; DB 10; Length 564;
Best Local Similarity 99.6%; Pred. No. 2.5e-86;
Matches 562; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 35 TGGTGTCTTCTCCCAACACATCTTGTCACGGTGGTGGCGGTTCATTTGGCCACCCACA 94
Db 1 TGGTGTCTTCTCCCAACACATCTTGTCACGGTGGTGGCGGTTCATTTGGCCACCCACA 60
Qy 95 CGTGTTCAGCTTCTCAAAAGCTGGCTTCAGGTTTCAATAATCGACAAATTCGATAACT 154
Db 61 CGTGTTCAGCTTCTCAAAAGCTGGCTTCAGGTTTCAATAATCGACAAATTCGATAACT 120
Qy 155 CGTTCATGGAAGCAGTGGACCGCGTCCGCAAGTGGTTGGCCCTCTGCTTCTCAGAACC 214
Db 121 CGTTCATGGAAGCAGTGGACCGCGTCCGCAAGTGGTTGGCCCTCTGCTTCTCAGAACC 180
Qy 215 TCCAAATTCACCCAGCGGATCTCCGGAATAGGGATGACTTCGAGAAACTCTTCTCCAAA 274
Db 181 TCCAAATTCACCCAGCGGATCTCCGGAATAGGGATGACTTCGAGAAACTCTTCTCCAAA 240
Qy 275 CAACATTTGATCGCGTGTATCCACTTCTGCTGCTTGAAGCGTTCGTAAGCGTTCGGA 334
Db 241 CAACATTTGATCGCGTGTATCCACTTCTGCTGCTTGAAGCGTTCGTAAGCGTTCGGA 300
Qy 335 AGCCCGTGGCTATTTTGAATTTTAAATTTGGTGGCACCATCAACCTCTACGAGTTTATGG 394
Db 301 AGCCCGTGGCTATTTTGAATTTTAAATTTGGTGGCACCATCAACCTCTACGAGTTTATGG 360
Qy 395 CAAGTATAATTCGAAAGATGTTTTCTCATCTGCAACCGTTTATGCGCAACCTG 454
Db 361 CAAGTATAATTCGAAAGATGTTTTCTCATCTGCAACCGTTTATGCGCAACCTG 420
Qy 455 AAAAGATACCGTGTGAGGAGGATTTCAAGTTTCAAGCTATGAATCCCTATGAGGAGCA 514
Db 421 AAAAGATACCGTGTGAGGAGGATTTCAAGTTTCAAGCTATGAATCCCTATGAGGAGCA 480
Qy 515 AGCTTTCTGGAAGAAATTCGCGAGATATTCGAAAGCTGAAACCGATGGAAGATCA 574
Db 481 AGCTTTCTGGAAGAAATTCGCGAGATATTCGAAAGCTGAAACCGATGGAAGATCA 540
Qy 575 TATTACTGAGATCTTCAATCCAG 598
Db 541 TATTACTGAGATCTTCAATCCAG 564

CB892692 798 bp mRNA linear EST 24-APR-2003
EST645484 HOGA Medicago truncatula cDNA clone HOGA-19B13, mRNA
sequence.
CB892692
CB892692.1 GI:30099860
EST.
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 798)
Hahn,M.G., Ojansen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
Utterback,T., Cho,J. and Frazer,C.M.
ESTs from roots of Medicago truncatula treated with
clisgagalacturonides of DP 6-20
Unpublished (2001)
JOURNAL

```

digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORt1 vector. The ligated cDNA fragments were transformed into E. coli ElectroMax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note that Gm-r1030 is a re-rack of Gm-cl007."

COMMENT

Contact: Michael G. Hahn
Complex Carbohydrate Research Center
University of Georgia
220 Riverbend Road, Athens, GA 30602-4712, USA
Tel: 706-542-4457
Fax: 706-542-4412
Email: hahn@ccrc.uga.edu
TIGR sequence name: MTMCW07TK
More information is available at: www.medicago.org
Seq primer: Skmcd (CTA gaa cta gta gaa cc).
Location/Qualifiers

FEATURES

source
1..798

/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="HOGA-19B13"
/tissue_type="3 day old seedling roots"
/dev_stage="24 hours after treatment in the dark at 26 C with 0.5 mg/ml oligogalacturonides (DP 6-20) in the presence of 100 ug/ml Gentamicin"
/lab_host="XLOLR"
/clone_lib="HOGA"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in SOLR cells."

ORIGIN

Query Match 42.7%; Score 560.2; DB 14; Length 798;
Best Local Similarity 83.3%; Pred. No. 3.4e-86;
Matches 637; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 13 TCCTCCTCTCTATTGCGACGATGGTCTCTCCACACATCTCGTCAACCGGTGT 72
DB 34 TCCTCCTCTCTATTGCGACGATGGTCTCTCCACACATCTCGTCAACCGGTGT 93
QY 73 GCGCGTTTCAATGGCCACCAACCGTGTTCAGTCTTCTCAAGCTGGTTCAGCGTTCA 132
DB 94 TCCTGTTTCAATGGCCACCAACCGTGTTCAGTCTTCTCAAGCTGGTTCAGCGTTCC 153
QY 133 ATATCGCAATTCGATTAATCTCGTTCATGAGCAGTGGACCGTTCGCGCAAGTGT 192
DB 154 ATTATTGATAATTCGATTAATCTCGTTCATGAGCAGTGGACCGTTCGCGCAAGTGT 213
QY 193 GCGCCTCTGTTTCTCAGAACCTCCCAATTCACCCAGCGGATCTCCGGAATAGGATGAC 252
DB 214 GGTCTCTCAATTCGATTAATCTCGTTCATGAGCAGTGGACCGTTCGCGCAAGTGT 273
QY 253 TTGAGAACTCTCTCCAAACCAACATTTGATGCGGTGATCCACTTTCGCGTTGAA 312
DB 274 TTGAGAACTCTCTCTAAACCAACATTTGATGCGGTGATCCACTTTCGCGTTGAA 333
QY 313 GCGGTTCTCAAGCGTTGCGAACCGCGTCTGATTTTGAATTTAATTTGGTTGCAAC 372
DB 334 GCGGTTGTTGAAGTTGCGAATCCCGTCTGATTTTGAATTTAATTTGGTTGCAAC 393
QY 373 ATCACTCTACAGTTTATGGCAAGTATTAATGCAAAAGATGTTTCTTCATCATCT 432
DB 394 ATCAATCTTATGAAGTTATGGCAAGTATTAATGCAAAAGATGTTTCTTCATCATCT 453
QY 433 GCACCGTTTATGGCAAGTATTAATGCAAAAGATGTTTCTTCATCATCT 492
DB 454 GCTACTGTTTATGGTCAACCTGACGATACCTGTTGGAGATTTCAAGTTACAGCC 513
QY 493 ATGAATCCCTATGGACGCAACAGCTTTTCTCGAAGAAATGCGCGAGATTTACAGAA 552
DB 514 ATGAATCCCTATGGCGGCACTAAGCTTTTCTTGAAGAAATGCGCGAGATTTACAGAA 573
QY 553 GCTGAACCAAGTGAAGATCATATTACTGAGATCTTCAATCCAGTGGGGTCAAGAA 612

DB 574 CTTGACCAACAGATGGAGATCATTTTACTGAGATCATTTTCAATCCAGTTGGAGCCCATGAA 633
QY 613 AGTGGCAAACTCGGTGAAGATCCCAAGGCGATCCCAAAATCACTCATGCTTACATCAG 672
DB 634 AGTGGTAAACTCGGTGAAGATCCCAAGGCGATCCCAAAATCACTCATGCTTACATCAG 693
QY 673 CAAGTAGCTGTTGGAAGATTTGACTGAACCTCAATGTATACGGTTCATGATTTATCCAAAGGAG 732
DB 694 CAAGTAGCTGTTGGAAGATTTGACTGAACCTCAATGTATATGGTTCATGATTTATCCACAGG 753
QY 733 GATGGCTCTCGGATCCGGACTATATCCATGTGATGACCTTGCGA 777
DB 754 GATGGCTCTCGGATCCGGACTATATCCATGTGATGACCTTAGCA 798

RESULT 10

AWS07982

LOCUS

DEFINITION

Gm-r1030-2084 5' similar to TR:Q43070 Q43070

UDP-GALACTOSE-4-EPIMERASE ; mRNA sequence.

ACCESSION

AW507982

AW507982.1 GI:7146060

VERSION

EST.

KEYWORDS

Glycine max (soybean)

SOURCE

Glycine max

ORGANISM

Public Soybean EST Project

REFERENCE

1. (bases 1 to 580)

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Corvelli, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE

Public Soybean EST Project

JOURNAL

Unpublished (1999)

COMMENT

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

Insert Length: 1230 Std Error: 0.00

High quality sequence stop: 458.

FEATURES

Location/Qualifiers

1..580

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clones="GENOME SYSTEMS CLONE ID: Gm-r1030-2084"

/lab_host="DH10B"

/clone_lib="Gm-r1030"

/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This

cDNA library was constructed from mRNA isolated from

immature cotyledons of greenhouse grown plants

(individual seed fresh weight of 100-300mg). The library

was prepared using the Life Technologies pSuperScript cDNA

library construction kit. Complementary DNA was

synthesized from mRNA using a poly(dT) sequence with a

NotI restriction site. SalI linkers followed by NotI

to the blunt-ended cDNA fragments were directionally cloned

digestion. The cDNA fragments were directionally cloned

into the NotI-SalI restriction site of the pSPORT1

vector. The ligated cDNA fragments were transformed into

E. coli ElectroMax DH10B host cells. This library was

constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note that Gm-1030 is a re-rack of Gm-cl007."

ORIGIN

Query Match 42.6%; Score 558.8; DB 10; Length 580;
Best Local Similarity 97.8%; Pred. No. 6.9e-86;
Matches 566; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 42 TTCTCTCCCAACACATTCGTGTCACCGGTGGTGGCGGTTTCATTTGGCACCCACACCGTGGT 101
DB 1 TTCTCTCCCAACACATTCGTGTCACCGGTGGTGGCGGTTTCATTTGGCACCCACACCGTGGT 60

QY 102 TCAGCTTCTCAAGCTGGTTCAGCGTTTCATATATCGACAAATTCGATATCCCGTCAT 161
DB 61 TCAGCTTCTCAAGCTGGTTCAGCGTTTCATATATCGACAAATTCGATATCCCGTCAT 120

QY 162 GGAAGCAGTGGACCGGTGGTGGCGGTGGTGGCGGTGGTGGCGGTGGTGGCGGTGGTGGCGGTGGT 221
DB 121 GGAAGCAGTGGACCGGTGGTGGCGGTGGTGGCGGTGGTGGCGGTGGTGGCGGTGGTGGCGGTGGT 180

QY 222 CACCCAGGGGATCTCCGGAATAGGATGACATTTGGAGAACTCTTCTCCAAACACACAT 281
DB 181 CACCCAGGGGATCTCCGGAATAGGATGACATTTGGAGAACTCTTCTCCAAACACACAT 240

QY 282 TGATGCGGTGATCCATTTGCTGCTTTGAAAGGGTGTGAAAGGGTGTGAAAGGGTGTGAAAGGGTGT 341
DB 241 TGATGCGGTGATCCATTTGCTGCTTTGAAAGGGTGTGAAAGGGTGTGAAAGGGTGTGAAAGGGTGT 300

QY 342 TCCTATTTTGAATTTAATTTGTTGGACCACTCAACCTCTACGAGTTATGGCAAGTA 401
DB 301 TCCTATTTTGAATTTAATTTGTTGGACCACTCAACCTCTACGAGTTATGGCAAGTA 360

QY 402 TAATTCGAAAGATGTTTTCATCATCTGCAACCGTTTATGGCCAACTGAAAGAT 461
DB 361 TAATTCGAAAGATGTTTTCATCATCTGCAACCGTTTATGGCCAACTGAAAGAT 420

QY 462 ACCGTGTGAGGAGATTTCAAGTTTCAAGCTATGATCCCTATGAGCGGACCAAGTTT 521
DB 421 ACCGTGTGAGGAGATTTCAAGTTTCAAGCTATGATCCCTATGAGCGGACCAAGTTT 480

QY 522 CTTGGAAGAAATTCGCGAGATATTCGAAAGCTGACCGAGATGGAAGATCATATTACT 581
DB 481 CTTGGAAGAAATTCGCGAGATATTCGAAAGCTGACCGAGATGGAAGATCATATTACT 540

QY 582 GAGATATTTCAATCCAGTTGGGGCTCATGAAAGTGCAA 620
DB 541 GAGATATTTCAATCCAGTTGGGGCTCATGATAGTGCAA 579

RESULT 11

Al856802
LOCUS sb78d02.v1 Gm-cl010 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl010-1084 5' similar to TR:Q43070 Q43070
USP-GALACTOSE-4-EPIMERASE ; mRNA sequence.

Al856802

Al856802.1 GI:5510418

VERSION EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

1 (bases 1 to 564)

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,I., Jackson,I., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project

Unpublished (1999)

REFERENCE

AUTHORS

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Possible reversed clone: similarity on wrong strand This clone is
available through: ResGen, Invitrogen Corp. 2130 South Memorial
Parkway Huntsville, AL 35801 For further information call:
(800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 411.

FEATURES

source

1..564
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl010-1084"
/tissue_type="young cotyledons of greenhouse grown plants"
/dev_stage="2cm long 12 week old"
/lab_host="XL10-Gold"
/clone_lib="Gm-cl010"
/note="vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature cotyledons (100-200mg) of old greenhouse
grown plants. The cDNA library was prepared using the
Stratagene pBluescript II SK(+) library construction kit.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly (dT) sequence with a XhoI restriction
site. EcoRI adapters were ligated to the blunt-ended cDNA
fragments followed by XhoI digestion. The cDNA fragments
were directionally cloned into the EcoRI-XhoI restriction
site of the pBluescript vector. The ligated cDNA fragments
were transformed into XL10-Gold host cells. This library
was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

ORIGIN

Query Match 42.6%; Score 558.4; DB 9; Length 564;
Best Local Similarity 99.6%; Pred. No. 8.1e-86;
Matches 559; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 CCATCTCTCCCTCTCTATTGACGATGTTCTCTCTCCCAACACATCTCGTCAACG 67
DB 2 CCATCTCTCTCCCTCTCTATTGACGATGTTCTCTCTCCCAACACATCTCGTCAACG 61

QY 68 GTGGTCCCGGTTTCATTGGCACCCACACCGTCTCAGCTTCTCAAAGTGGCTTCAGCG 127
DB 62 GTGGTCCCGGTTTCATTGGCACCCACACCGTCTCAGCTTCTCAAAGTGGCTTCAGCG 121

QY 128 TTTCAATATCGACAATTCGATTAATTCCTGATGAGCAGTGCAGCGTCCGCCAAG 187
DB 122 TTTCAATATCGACAATTCGATTAATTCCTGATGAGCAGTGCAGCGTCCGCCAAG 181

QY 188 TGGTGGCCCTCTGCTTTCTCAGAACCTTCAATTCACCCAGGGCGATCTCCGGAATAGG 247
DB 182 TGGTGGCCCTCTGCTTTCTCAGAACCTTCAATTCACCCAGGGCGATCTCCGGAATAGG 241

QY 248 ATGACTTGGAGAACTCTTCTCCAAACACATTTGATGCGGTGATCCACATTTGCTGGCT 307
DB 242 ATGACTTGGAGAACTCTTCTCCAAACACATTTGATGCGGTGATCCACATTTGCTGGCT 301

QY 308 TGAAGACGGTGTGCGAAGCGTTGCGAAGCCCGCTCGCTATTTTGAATTTAATTTGGTTG 367
DB 302 TGAAGACGGTGTGCGAAGCGTTGCGAAGCCCGCTCGCTATTTTGAATTTAATTTGGTTG 361

QY 368 GCACCATCAACTCTACAGTTTATGGCAAGATATTAATTCGAAAGATAGTGTTCAT 427
DB 362 GCACCATCAACTCTACAGTTTATGGCAAGATATTAATTCGAAAGATAGTGTTCAT 421

QY 428 CATCTGCACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGGAGATTTCAAGTTAC 487
DB 422 CATCTGCACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGGAGATTTCAAGTTAC 481

South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: cu@resgen.com web site:

www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 435.

Location/Qualifiers

FEATURES

source

1. .547
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl061-3235"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl061"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 41.2%; Score 540.6; DB 12; Length 547;
Best Local Similarity 99.3%; Pred. No. 9.1e-83;
Matches 543; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 627 TGAAGATCCCAAGGCGATCCCAATACCTTCATGCTTACCTTACATTCAGCAAGTAGCTGTGG 686
Db 1 TGAAGATCCCAAGGCGATCCCAATACCTTCATGCTTACATTCAGCAAGTAGCTGTGG 60
Qy 687 AAGATTGACCTGAATCAATGATACGGTCTATGATATTCACACGAGGATGGCTCTCGAT 746
Db 61 AAGATTGACCTGAATCAATGATACGGTCTATGATATTCACACGAGGATGGCTCTCGAT 120
Qy 747 CCGGACTATATCCATGTGATGACCTTGGCAGATGCCATATTGCTGCCCTCGAAAGCT 806
Db 121 CCGGACTATATCCATGTGATGACCTTGGCAGATGCCATATTGCTGCCCTCGAAAGCT 180
Qy 807 CTTTCACACGAGACATAGTGTGACTGCTTACACCTGGGACCTGCTGGACATC 866
Db 181 CTTTCACACGAGACATAGTGTGACTGCTTACACCTGGGACCTGCTGGACATC 240
Qy 867 TGTGCTTGAATGGTTACAGCAATTTGAAAAGGCTTCTGGCAAGAAATTCAGTAAATTT 926
Db 241 TGTGCTTGAATGGTTACAGCAATTTGAAAAGGCTTCTGGCAAGAAATTCAGTAAATTT 300
Qy 927 ATGTCGAAGAACCGGAGATCGGACTGAGTTTATGATCTACAGAGAGCTGAGAA 986
Db 301 ATGTCGAAGAACCGGAGATCGGACTGAGTTTATGATCTACAGAGAGCTGAGAA 360
Qy 987 AGAATCTGTTGGAAGCAAACTATGTTGTGGAGGAGATGTGCAGCGACCAATGGAAATTG 1046
Db 361 ATAACTTGTGTGAGAGCAAACTATGTTGTGGAGGAGATGTGCAGCGACCAATGGAAATTG 420
Qy 1047 GGCAGAGAAACAACTCCCTGGGGTTACCGCGGAGAGCCCTTGAATGCTTGAAGAAATATCT 1106
Db 421 GGCAGAGAAACAACTCCCTGGGGTTACCGCGGAGAGCCCTTGAATGCTTGAAGAAATATCT 480
Qy 1107 GCTCATCTAGAAATGCTTTTACATAAATAGGATCTCTTATATAGATACATTTTATGTT 1166
Db 481 GCTCATCTAGAAATGCTTTTACATAAATAGGATCTCTTATATAGATACATTTTATGTT 540
Qy 1167 TGATGAT 1173
Db 541 GGATGAT 547

RESULT 14

BF635019

LOCUS

DEFINITION

BF635019

ACCESSION

BF635019

KEYWORDS

SOURCE

ORGANISM

Medicago truncatula (barrel medic)

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

1 (bases 1 to 694)

REFERENCE

AUTHORS

Flores-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,

Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula drought library

Unpublished (2000)

JOURNAL

COMMENT

Contact: May GD

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: gdmay@noble.org

Insert Length: 694 Std Error: 0.00

Plate: 078 row: G column: 09

Seq primer: TCACACAGGAACACAGCTATGAC.

Location/Qualifiers

1. 694

/organism="Medicago truncatula"

/mol_type="mRNA"

/db_xref="taxon:3880"

/clone="NF07809DT"

/tissue_type="Plantlets"

/dev_stage="Pooled timepoints"

/clone_lib="Drought"

/note="Vector: Lambda Zap; Contains a mixture of entire

plantlets harvested in a series of days-post-watering

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

timepoints."

GenCore version 5.1.6
Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 20:18:05 ; Search time 74 Seconds
(without alignments)
1336.373 Million cell updates/sec

Title: US-09-913-064A-14

Perfect score: 1845

Sequence: 1 MVSSSQHLVTGAGFIGTH.....MCRDQWNAKNPWYAGKP 350

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1845	100.0	350	3	AAB07819 Amino aci
2	1530	82.9	351	3	AAG09987 Arabidops
3	1486	80.5	351	5	AAG79852 Potato DG
4	1384	75.0	352	3	AAB07820 Amino aci
5	1358	73.6	339	3	AAG15804 Arabidops
6	1231	66.7	375	3	AAG15222 Arabidops
7	1231	66.7	377	3	AAG42840 Arabidops
8	1229	66.6	348	3	AAG15224 Arabidops
9	1229	66.6	349	3	AAG15223 Arabidops
10	1229	66.6	350	3	AAG42842 Arabidops
11	1229	66.6	351	3	AAG42841 Arabidops
12	1222	66.2	348	3	AAG53259 Arabidops
13	1222	66.2	372	3	AAG53258 Arabidops
14	1220	66.1	348	3	AAG321995 Arabidops
15	1220	66.1	372	3	AAG321994 Arabidops
16	1208	65.5	350	3	AAG51824 Arabidops
17	1201	65.1	350	3	AAG42204 Arabidops
18	1192	64.6	349	3	AAB07823 Amino aci
19	1180	64.0	354	3	AAB07822 Amino aci
20	1094	59.3	231	3	AAG09988 Arabidops
21	1054	57.2	224	3	AAB07821 Amino aci
22	1054	57.1	224	3	AAG09989 Arabidops
23	1038	56.3	339	4	AAG64602 Uridine d
24	1038	56.3	339	5	AB881469 B. subtil
25	1012.5	54.9	266	3	AAG53260 Arabidops

26	1010.5	54.8	266	3	AAG21996	Aag21996 Arabidops
27	986.5	53.5	348	3	AAV54118	Aay54118 Amino aci
28	986.5	53.5	348	5	ABG96567	Human sho
29	977.5	53.0	348	2	AAW01619	Human ur1
30	977.5	53.0	348	4	AAV51152	Human ur1
31	977.5	53.0	348	7	ADA66613	Human ur1
32	966	52.4	339	6	ABU02070	S. pneumo
33	952	51.6	341	6	ABM68787	Phototrab
34	941	51.0	336	6	ABU02313	S. pneumo
35	927	50.2	340	5	ABP66248	Bifidobac
36	918.5	49.8	350	4	ABBS7926	Drosophil
37	917.5	49.7	341	6	ADB09548	Alloiooc
38	904.5	49.0	338	5	ABB81470	N. gonorr
39	904.5	49.0	340	6	ABP80860	N. gonorr
40	904.5	49.0	340	6	ABP80439	N. gonorr
41	898	48.7	219	3	AAG15805	Arabidops
42	893.5	48.4	338	3	AAV54117	Amino aci
43	891.5	48.3	338	3	AAV81359	Escherich
44	887.5	48.1	235	3	AAG51825	Arabidops
45	878.5	47.6	699	5	ABG93041	S. cerevi

ALIGNMENTS

RESULT 1

AAB07819
ID AAB07819 standard; protein; 350 AA.

AC AAB07819;

DT 14-NOV-2000 (first entry)

DE Amino acid sequence of UDP-galactose 4-epimerase.

XX UDP-galactose 4-epimerase; UDP-glucose; carbon flux;
KW raffinose pathway; raffinose; stachyose; carbohydrate metabolism;
KW starch level.

OS Glycine max.

XX Key Location/Qualifiers

FT Misc-difference 3...129
FT /notes="the nucleotides encoding these residues are not given"

XX WO200047755-A2.

XX 17-AUG-2000.

XX 09-FEB-2000; 2000WO-US003453.

XX 10-FEB-1999; 99US-0119588P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PI Carlson TJ, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;

PI Rafalski JA, Thorpe CJ;

XX WPI; 2000-5495152/50.

XX N-ESDB; AAA59439.

XX Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant cells to e.g. decrease cell wall constituents and increase starch content.

XX Claim 10; Page 51-52; 61pp; English.

XX The present sequence represents a UDP-galactose 4-epimerase. The enzyme is responsible for the interconversion of UDP-glucose and UDP-galactose. As the activity of the enzyme appears to be particularly limiting to carbon flux into the raffinose pathway, reduction of its activity should decrease the levels of raffinose and stachyose in seeds. The


```

PR 01-FEB-2001; 2001US-0265311P.
XX (DAIZ/) DAI Z.
XX
XX Dai Z, Shi L, Hooker BS;
XX
XX WPI; 2002-698747/75.
XX N-PSDB; ABA00335.
XX
XX New isolated polynucleotide molecule comprising a gene coding for the
XX enzyme 5'-diphospho galactose 4-epimerase, useful for regulating
XX carbohydrate metabolism or nutritional profile in transgenic plants.
XX
XX Claim 1; Fig 1; 38pp; English.
XX
XX This sequence is encoded by the potato psen-1 gene and represents the
XX enzyme 5'-diphospho-galactose 4-epimerase (DGE). This sequence is
XX homologous to the sequence of Arabidopsis thaliana uridine diphosphate
XX glucose (UDPG) epimerase gene (86% homology at peptide level, 71%
XX homology at the nucleotide level). The DGE polynucleotide molecules are
XX useful in regulating carbohydrate metabolism in transgenic plants. The
XX polynucleotide can be used in Controlled Environment Agriculture that
XX employs an integrated system for commercial production of transgenic
XX plants in a controlled environment. Controlling the level of DGE in
XX transgenic plants is useful for regulating the nutritional profile of the
XX plant
XX
XX Sequence 351 AA;
XX
XX Query Match 80.5%; Score 1486; DB 5; Length 351;
XX Best Local Similarity 78.2%; Pred. No. 3.3e-133;
XX Matches 273; Conservative 33; Mismatches 43; Indels 0; Gaps 0;
XX
XX QY 2 VSSSOHLVTGGAGTGTHTVOLLKAGFSVSIIDNFNSVMEAVDRVRQVVGPLLSQNL 61
XX Db 3 VQCEMLVTGGAGTGTHTVOLLNEGPKVTIIDNFNSVEAVDRVRELVPQLSQNL 62
XX
XX QY 62 QFTQGLRRDDLEKLFSTTDFDAVTHFAGLKAVAESVAKPRYPDFNLVGTINLYEFMA 121
XX Db 63 EPHLGDIRKDDLEKLFKSKEFAAVVHFAGLKAVGESVQVPELYFENNLSITLYSMA 122
XX
XX QY 122 KYNCKMWPSSSATVYGOPEKIPCEEDFKLQAMNPGYGRTKLFLFEEIARDIOKAEPEWKII 181
XX Db 123 KYNCKLVSSTATVYGOPEKIPCEVEDFKAMNPGYGRTKLFLFEEIARDIOKADQEWNI 182
XX
XX QY 182 LLRYFNPVGAHESGKLGEDPKGIPNNLMFYIQQAVAGRLTELVNHYGHDYPTDGSAIRDY 241
XX Db 183 LLRYFNPVGAHESGKLGEDPKGIPNNLMFYIQQAVAGRLPELVNHYGNDYPTDGTAIRDY 242
XX
XX QY 242 IHVMDLADGHIAALRKLFTEENIGCTAYNLGRTGTSVLEMTAPEKAGKKIPVKLCPR 301
XX Db 243 IHVLDLADGHVVALQRLLRQNLHGCVAYNLGRTGKSKSVLEMTAAFERASGKKIPLKMCPR 302
XX
XX QY 302 RPDGATEVYASTERAEKELGWKANYGVEMCRDQMWAKNPMGYAGKP 350
XX Db 303 RPDGATVYASTEAEKELGWAKYGINEMCRDQMWAKNSQNPWGYSKP 351
XX
XX RESULT 4
XX AAB07820
XX ID AAB07820 standard; protein; 352 AA.
XX
XX AC AAB07820;
XX
XX DT 14-NOV-2000 (first entry)
XX
XX Amino acid sequence of UDP-galactose 4-epimerase.
XX
XX UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
XX raffinoseaccharide pathway; raffinose; stachyose; carbohydrate metabolism;
XX starch level.
XX
XX Triticum aestivum.
XX
XX 01-FEB-2001; 2000WO-US003453.
XX 17-AUG-2000.
XX
XX 09-FEB-2000; 2000WO-US003453.
XX
XX 10-FEB-1999; 99US-0119588P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Carlson TJ, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;
XX Rafalski JA, Thorpe CJ;
XX
XX WPI; 2000-549152/50.
XX N-PSDB; AAA59440.
XX
XX Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant
XX cells to e.g. decrease cell wall constituents and increase starch
XX content.
XX
XX Claim 10; Page 52-53; 61pp; English.
XX
XX The present sequence represents a UDP-galactose 4-epimerase. The enzyme
XX is responsible for the interconversion of UDP-glucose and UDP-galactose.
XX As the activity of the enzyme appears to be particularly limiting to
XX carbon flux into the raffinoseaccharide pathway, reduction of its activity
XX should decrease the levels of raffinose and stachyose in seeds. The
XX polynucleotide is useful to transform plant cells to enhance the degree
XX of expression of UDP-galactose-4-epimerase within these cells. Changes in
XX the expression of UDP-galactose-4-epimerase within a cell enables
XX modification of plant carbohydrate metabolism. This allows growth of
XX grains with reduced cell wall constituents (fiber) and increased levels
XX of starch
XX
XX Sequence 352 AA;
XX
XX Query Match 75.0%; Score 1384; DB 3; Length 352;
XX Best Local Similarity 73.6%; Pred. No. 1.9e-123;
XX Matches 251; Conservative 42; Mismatches 48; Indels 0; Gaps 0;
XX
XX QY 8 ILVTGGAGTGTHTVOLLKAGFSVSIIDNFNSVMEAVDRVRQVVGPLLSQNLQFTQGD 67
XX Db 5 VLVTGGAGTGTHTVOLLLEKGYAVTAVDNFNSVPEALDRVRHIVGPALSARLQIFGD 64
XX
XX QY 68 LRNRDDLEKLFSTTDFDAVTHFAGLKAVAESVAKPRYPDFNLVGTINLYEFMAKYNCKK 127
XX Db 65 LTIKDDLEKVFAAKKYDAVTHFAGLKAVAESVAHPENYRNNIVGTNLYDMKGGCNK 124
XX
XX QY 128 MVFSSSATVYGOPEKIPCEEDFKLQAMNPGYGRTKLFLFEEIARDIOKAEPEWKIILLRYFN 187
XX Db 125 LVFSSSATVYGOPEKIPCEEDFKLQAMNPGYGRTKLFLFEEIARDIOKAEPEWKIILLRYFN 184
XX
XX QY 188 PVGAHESGKLGEDPKGIPNNLMFYIQQAVAGRLTELVNHYGHDYPTDGSAIRDYIHVMDL 247
XX Db 185 PIGAHESGDIGEDPKGIPNNLLPYIQQAVARRPELVNHYGHDYTRDTGTAVDYIHVVDL 244
XX
XX QY 248 ADGHIAALRKLFTEENIGCTAYNLGRTGTSVLEMTAPEKAGKKIPVKLCRRRPGDAT 307
XX Db 245 ADGHIAALEKLFATPDIGCVAYNLGRTGTSVLEMTAPEKAGKKIPVKLCRRRPGDSE 304
XX
XX QY 308 EVYASTERAEKELGWKANYGVEMCRDQMWAKNPMGYAG 348
XX Db 305 QVYASTAKAEELGWRKAKYGIEMCRDQMWAKNPMGYGCG 345
XX
XX RESULT 5
XX AAG15804
XX ID AAG15804 standard; protein; 339 AA.
XX
XX AC AAG15804;
XX
XX DT 17-OCT-2000 (first entry)

```

XX Arabidopsis thaliana protein fragment SEQ ID NO: 16201.
DE XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129645P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140659P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144632P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145088P.
XX 22-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
XX 22-JUL-1999; 99US-0145089P.
XX 22-JUL-1999; 99US-0145152P.
XX 23-JUL-1999; 99US-0145145P.
XX 23-JUL-1999; 99US-0145218P.
XX 26-JUL-1999; 99US-0145224P.
XX 27-JUL-1999; 99US-0145913P.
XX 27-JUL-1999; 99US-0145918P.
XX 27-JUL-1999; 99US-0145919P.
XX 28-JUL-1999; 99US-0145951P.
XX 02-AUG-1999; 99US-0146386P.
XX 02-AUG-1999; 99US-0146388P.
XX 02-AUG-1999; 99US-0146389P.
XX 03-AUG-1999; 99US-0147038P.
XX 04-AUG-1999; 99US-0147204P.
XX 04-AUG-1999; 99US-0147302P.
XX 05-AUG-1999; 99US-0147192P.
XX 05-AUG-1999; 99US-0147260P.
XX 06-AUG-1999; 99US-0147303P.
XX 06-AUG-1999; 99US-0147416P.
XX 09-AUG-1999; 99US-0147492P.
XX 09-AUG-1999; 99US-0147935P.
XX 10-AUG-1999; 99US-0148171P.
XX 11-AUG-1999; 99US-0148319P.
XX 12-AUG-1999; 99US-0148341P.
XX 13-AUG-1999; 99US-0148565P.
XX 13-AUG-1999; 99US-0148684P.
XX 16-AUG-1999; 99US-0149368P.
XX 17-AUG-1999; 99US-0149175P.
XX 18-AUG-1999; 99US-0149426P.
XX 20-AUG-1999; 99US-0149722P.
XX 20-AUG-1999; 99US-0149723P.
XX 20-AUG-1999; 99US-0149902P.
XX 23-AUG-1999; 99US-0149903P.
XX 25-AUG-1999; 99US-0150566P.
XX 26-AUG-1999; 99US-0150884P.
XX 27-AUG-1999; 99US-0151065P.
XX 27-AUG-1999; 99US-0151066P.

```
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 08-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162143P.

Query Match 73.6%; Score 1358; DB 3; Length 339;
Best Local Similarity 85.3%; Pred. No. 5.5e-121;
Matches 259; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

QY 3 SSSQHILVTGGAGFIGHTHTVQLLKAGFSVSIIDNFDNSVMEADRVVRVGVGFLLSQNLQ 62
Db 4 SVEQNIILVTGGAGFIGHTHTVQLLKAGFSVSIIDNFDNSVIEADRVRELVLGFLSKLD 63
QY 63 FTGDLNRDDELEKFSKTTFDVAIHAGIKAVAESVAKPRRYDFDNLTGTLNLYEFMAK 122
Db 64 FNLGDLNRKGDIEKLFSKQRFDAVHIFAGIKAVGESVENPRRYFDNLTGTLNLYETMAK 123
QY 123 YNCKMVFSSSATVYQGEKIPCEEDFKLOANPYGRTKLFLEETIARDIQKAEPEWKIL 182
Db 124 YNCKMVFSSSATVYQGEKIPCEEDFKLOANPYGRTKLFLEETIARDIQKAEPEWKIL 183
QY 183 LRVFNVPVGHESGKLGEDPKGIENLMPYITQAVAGRLTENVYGHYTRDGSARDVI 242
Db 184 LRVFNVPVGHESGKLGEDPKGIENLMPYITQAVAGRLTENVYGHYTRDGSARDVI 243
QY 243 HVMDLADGHIALLRKLFITENIGCTAYNLGTGRGTSVLEMTAFERKAGKKIPVKLCPR 302
Db 244 HVMDLADGHIALLRKLFITENIGCTAYNLGTGRGTSVLEMTAFERKAGKKIPVKLCPR 303

OY 303 PGD 305
Db 304 SGD 306

RESULT 6
AAG15222
ID AAG15222 standard; protein; 375 AA.
AC AAG15222;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 15390.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
```


Db 85 SFHQVLDRLDRSALEKIFSETKFDVAIHFAGLKAVGESVEKPLLYNNLVGTTITLLEVMA 144
QY 122 KYNCKQWVFSSSATVYGOEKIPCEBDFKLQAMNPVGRTKLFLEBIARDIOKAEPEWKII 181
Db 145 QHCCKNLVFSSTATVYGSKEVPCTEEFPIISALNPYGRTKLFIEICRDVYGSDFEWKII 204
QY 182 LLRYFNPVGAHSGKLGEDPKGIPNNLMPIYIQQAVGRITELNLYGHDPTRDGSAIRDY 241
Db 205 LLRYFNPVGAHPSGDTGEDPRGIPNNLMPFVQQAVGRPHLTVEGNDYNTKDGTVGRDY 264
QY 242 IHVMDLADGHIALRLKLTFTENIGCTAYNLGTGRGTVLEMTAPEKASGKKIPVKLCPR 301
Db 265 IHVIDLADGHIALRLKLECK-IGCEVNLGTGNGTSVLEMTDAPEKASGKKIPLVIAGR 323
QY 302 RPDGATEVYASTERAKEIGLWKANYGVHEMCRDQNWAKNNPWGY 346
Db 324 RPDGAEVYASTERASSELNWKAKYIGIEBMCRLMNASNNPYG 368

RESULT 7
AAG42840
ID AAG42840 standard; protein; 377 AA.
XX AC AAG42840;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 53475.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KM termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127452P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139450P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145244P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.

PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144484P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147322P.
PR 05-AUG-1999; 99US-0147360P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

PR	21-OCT-1999;	99US-0160814P.	PR	01-APR-1999;	99US-0127462P.
PR	21-OCT-1999;	99US-0160815P.	PR	06-APR-1999;	99US-0128234P.
PR	22-OCT-1999;	99US-0160808P.	PR	08-APR-1999;	99US-0128714P.
PR	22-OCT-1999;	99US-01608081P.	PR	16-APR-1999;	99US-0129845P.
PR	22-OCT-1999;	99US-01608099P.	PR	19-APR-1999;	99US-0130077P.
PR	25-OCT-1999;	99US-0161404P.	PR	21-APR-1999;	99US-0130449P.
PR	25-OCT-1999;	99US-0161405P.	PR	23-APR-1999;	99US-0130510P.
PR	25-OCT-1999;	99US-0161406P.	PR	23-APR-1999;	99US-0130891P.
PR	26-OCT-1999;	99US-0161359P.	PR	28-APR-1999;	99US-0131449P.
PR	26-OCT-1999;	99US-0161360P.	PR	30-APR-1999;	99US-0132048P.
PR	26-OCT-1999;	99US-0161361P.	PR	30-APR-1999;	99US-0132407P.
PR	28-OCT-1999;	99US-0161920P.	PR	04-MAY-1999;	99US-0132484P.
PR	28-OCT-1999;	99US-0161952P.	PR	05-MAY-1999;	99US-0132485P.
PR	28-OCT-1999;	99US-0161993P.	PR	06-MAY-1999;	99US-0132486P.
PR	29-OCT-1999;	99US-0162142P.	PR	06-MAY-1999;	99US-0132487P.
Query Match 66.6%; Score 1229; DB 3; Length 348;			PR	07-MAY-1999;	99US-0132863P.
Best Local Similarity 66.7%; Pred. No. 1.2e-108;			PR	11-MAY-1999;	99US-0134256P.
Matches 226; Conservative 50; Mismatches 62; Indels 2; Gaps 2;			PR	14-MAY-1999;	99US-0134218P.
			PR	14-MAY-1999;	99US-0134219P.
Qy	5	SOHILVTGGAGFIGTHTVOLLKAGFSVSIIDNFDNSWMEAVDRVQVVGPLLSONLQFT 64	PR	14-MAY-1999;	99US-0134221P.
Db	2	ARNVLVSGGAGYIGSHVTQLLLGGYSVVVDNLNDSAVSLQKVKKLAAE-HGERLSEH 60	PR	14-MAY-1999;	99US-0134370P.
Qy	65	QGDLNRDDLEKLFSTTFDAVTHFAGLKAVASVAKPRYFDNVLGVTINLYEFVAKYN 124	PR	18-MAY-1999;	99US-0134768P.
Db	61	QVDLRDSALEKIFSETKFDAVTHFAGLKAVGSVEKPLLYYNNLVGTTLLVMAQHG 120	PR	19-MAY-1999;	99US-0134941P.
Qy	125	CKMVFSSATVYQPEKIPCEEDFKLOAMNPYGRTKLFLFEEIARDIQKAEPEWKILLR 184	PR	20-MAY-1999;	99US-0135124P.
Db	121	CKNLVFSSATVYQSPKEVCTEFPISALNPYGRTKLFFEEICRDVYSGDPEWKILLR 180	PR	21-MAY-1999;	99US-0135353P.
Qy	185	YFNPVGAHESKLGEDPKGIPNNIMPYIQQVAVGRLTTELNVYGHDPTRDGSAIRDIYHV 244	PR	21-MAY-1999;	99US-0135629P.
Db	181	YFNPVGAHPSGDI GEDPRGIPNNLMPFVQVAVGRREPHLTVFGNDYNTKDGTVGRDIYHV 240	PR	25-MAY-1999;	99US-0136021P.
Qy	245	MDLADGHIAALRLFTTENGICTAYNLGTGRGTVLEMTAFKASGKKIPVKLCRRPG 304	PR	27-MAY-1999;	99US-0136392P.
Db	241	IDLADGHIAALRKLEDCX-IGCEVYNLTGNGTSVLEMDAFKASGKKIPVIAGRRPG 299	PR	28-MAY-1999;	99US-0136782P.
Qy	305	DATEVYASTERAELGWKANYGVEMCRDOWNWAKNPNYGY 346	PR	01-JUN-1999;	99US-0137222P.
Db	300	DAEVYASTERAESLNWKAKYGIEMCRDLNWNASNNPYGY 341	PR	03-JUN-1999;	99US-0137528P.
RESULT 9			PR	04-JUN-1999;	99US-0137502P.
ID	AAG15223	standard; protein; 349 AA.	PR	07-JUN-1999;	99US-0137724P.
XX	AC	AAG15223;	PR	08-JUN-1999;	99US-0138094P.
XX	DT	17-OCT-2000 (first entry)	PR	10-JUN-1999;	99US-0138540P.
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 15391.	PR	10-JUN-1999;	99US-0138847P.
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;	PR	14-JUN-1999;	99US-0139119P.
XX	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	PR	16-JUN-1999;	99US-0139452P.
XX	KW	termination sequence.	PR	16-JUN-1999;	99US-0139453P.
XX	OS	Arabidopsis thaliana.	PR	17-JUN-1999;	99US-0139429P.
XX	PN	EP1033405-A2.	PR	18-JUN-1999;	99US-0139454P.
XX	PD	06-SEP-2000.	PR	18-JUN-1999;	99US-0139455P.
XX	PF	25-FEB-2000; 2000EP-00301439.	PR	18-JUN-1999;	99US-0139461P.
XX	PR	25-FEB-1999; 99US-0121825P.	PR	18-JUN-1999;	99US-0139462P.
PR	05-MAR-1999;	99US-0123180P.	PR	18-JUN-1999;	99US-0139463P.
PR	09-MAR-1999;	99US-0123548P.	PR	18-JUN-1999;	99US-0139750P.
PR	23-MAR-1999;	99US-0125788P.	PR	21-JUN-1999;	99US-0139763P.
PR	25-MAR-1999;	99US-0126264P.	PR	22-JUN-1999;	99US-0139817P.
PR	29-MAR-1999;	99US-0126785P.	PR	22-JUN-1999;	99US-0139899P.
PR			PR	23-JUN-1999;	99US-0140353P.
PR			PR	23-JUN-1999;	99US-0140354P.
PR			PR	24-JUN-1999;	99US-0140695P.
PR			PR	28-JUN-1999;	99US-0140823P.
PR			PR	29-JUN-1999;	99US-0140991P.
PR			PR	30-JUN-1999;	99US-0140991P.
PR			PR	01-JUL-1999;	99US-0141287P.
PR			PR	01-JUL-1999;	99US-0141842P.
PR			PR	02-JUL-1999;	99US-0142154P.
PR			PR	02-JUL-1999;	99US-0142055P.
PR			PR	06-JUL-1999;	99US-0142390P.
PR			PR	08-JUL-1999;	99US-0142803P.
PR			PR	09-JUL-1999;	99US-0142920P.
PR			PR	12-JUL-1999;	99US-0142977P.
PR			PR	13-JUL-1999;	99US-0143542P.
PR			PR	14-JUL-1999;	99US-0143624P.
PR			PR	15-JUL-1999;	99US-0144005P.
PR			PR	16-JUL-1999;	99US-0144085P.
PR			PR	16-JUL-1999;	99US-0144086P.
PR			PR	19-JUL-1999;	99US-0144325P.

PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	04-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	06-AUG-1999;	99US-0147493P.			
PR	08-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	12-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	16-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149923P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157153P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
<hr/>					
PR	13-OCT-1999;	99US-0159293P.	PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.	PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.	PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.	PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.	PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.	PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.	PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.	PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.	PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.	PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.	PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.	PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.	PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.	PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.	PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.	PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.	PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.	PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.	PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.	PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.	PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.	PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.	PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.	PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.	PR	29-OCT-1999;	99US-0162142P.
<hr/>					
Query Match 66.6%; Score 1229; DB 3; Length 349;					
Best Local Similarity 66.7%; Pred. No. 1.2e-108;					
Matches 228; Conservative 50; Mismatches 62; Indels 2; Gaps 2;					
QY	5	SOHLVTCGAGFCTHTVQLLKAGPSVSIIDNPDNSVMEAVDRVGVGPELLSONLQFT	64		
DB	3	ARNVLVGGAGYIGSHTVLQLLGGYSVVVDNLNDSASVLSQVKKLAAB-HGERLSFH	61		
QY	65	QGDLRNDDLEKLFSTTFDAVHFAGLKAVAESVAKPRYFDFNLVGTINLYEFMAKYN	124		
DB	62	QVDLRDRSALEKIFSEIKFDVAFHFAGLKAVGESVEKPLLYNNLVGTITLLEVMAQH	121		
QY	125	CKRVFSSATVYQPEKIPCEEDFKLOANPYGRTKLFLEETARDQKAPEWKILLR	184		
DB	122	CKNLVSSATVYSPKEVCTEFPISALNPYGRTKLFEEICRDVYSGDPEWKILLR	181		
QY	185	YFNPVGAHESKLGEDPKGHPNNLMPYIQVAVGRLTELNVYGHDPYTRDSAIRDYHV	244		
DB	182	YFNPVGAHPSODIGEDPRGIPNNLMPFVQVAVGRRPHLVFGNDYNTKDGTVRDIHV	241		
QY	245	MDLADGHIAALRKLTFTENIGCTAYNLGTRGTSVLEMTAFKASGKKIPVKLCPRRP	304		
DB	242	IDLADGHIAALRKLEDC-K-IGCEVYNLTGNGTSVLEMTAFKASGKKIPLVIAGRRP	300		
QY	305	DATEVYASTEAELGKWKANYGVVEEMCRDQWAKNPNWGY	346		
DB	301	DAEVYASTEAELGKWKANYGVVEEMCRDQWAKNPNWGY	342		
<hr/>					
RESULT 10					
AAG42842					
ID	AAG42842	standard; protein; 350 AA.			
XX	AC	AAG42842;			
XX	DT	19-OCT-2000 (first entry)			
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 53477.			
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;			
XX	KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
XX	OS	termination sequence.			
XX	OS	Arabidopsis thaliana.			

XX EPI033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126284P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139753P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.

PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	28-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158322P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161932P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 66.4%; Score 1229; DB 3; Length 350;			
Best Local Similarity 66.7%; Pred. No. 1.2e-108;			
Matches 228; Conservative 50; Mismatches 62; Indels 2; Gaps 2;			
Qy	5	SOHLVGGAGFGTHTVWOLLKAGFSVSIIDNFDSVMEAVDRVQVYVGPILLSQNLQFT 64	
Db	2	ARNVLVGGAGYIGSHTVLQLLGGYSVVVDNLDNSAVSQRVKKLAAE-HGERLSPH 60	
Qy	65	QGLRNRDDLEKLFSTTFDAVHFAGLKAVASVAKPRRYFDNVLVGTINLYEFNAKYN 124	
Db	61	QVDLRDSALEKIFSETKFDVHFAGLKAAGSVSEKPLLYNNLVGTTLLLEVMAQHG 120	
Qy	125	CKMVFSSSTVYQQEPIKCEBDFKLQANPYGRKLFLEETARDIQKAEPEWKILLR 184	
Db	121	CKNLVFSSATVYGSPEVCTEPTFISALNPYGRKLFLEETCRDVGSDPEWKILLR 180	
Qy	185	YFNPVGAHESGKLGDEPKGTPNNLMPYIQOVAVGRLTGELNVYGHDPYTRDGSAIRDIHV 244	
Db	181	YFNPVGAHPGSDIGEDPRGTPNNLMPYIQOVAVGRPHLTGNDYNTKGTGVRDIHV 240	
Qy	245	MDLADGHIARLKRFTTENTIGCTAYNLGTGRGTSVLEMTVAFKAGKKIPVKLCPRRRG 304	
Db	241	IDLADGHIARLKRLEDCCK-IGCEVYNLGTNGTGVLEWDAFEKAGSKKIPLVIAGRRPG 299	
Qy	305	DATEVYASTERAEKLGKANYGVENECRDWNWAKNNPWGY 346	
Db	300	DAEVYASTERASELKNWAKYGLEEMCRDLNWNASNNPYG 341	

ID	AAG42841 standard; protein; 351 AA.
XX	AAG42841;
AC	
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 53476.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127482P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	04-MAY-1999; 99US-0132407P.
PR	05-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	14-MAY-1999; 99US-0134370P.
PR	18-MAY-1999; 99US-0134768P.
PR	19-MAY-1999; 99US-0134941P.
PR	20-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137282P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
PR	08-JUN-1999; 99US-0138094P.
PR	10-JUN-1999; 99US-0138540P.
PR	10-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
PR	18-JUN-1999; 99US-0139454P.
PR	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.
PR	18-JUN-1999; 99US-0139457P.
PR	18-JUN-1999; 99US-0139458P.
PR	18-JUN-1999; 99US-0139459P.
PR	18-JUN-1999; 99US-0139460P.

```
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142350P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143644P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145132P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145931P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147182P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.

PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 03-OCT-1999; 99US-0157117P.
PR 03-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.6%; Score 1229; DB 3; Length 351;
Best Local Similarity 66.7%; Pred. No. 1.3e-108;
Matches 228; Conservative 50; Mismatches 62; Indels 2; Gaps 2;

QY 5 SOHILVTGGAGFIGTHTVVQLLKAGFSVSIIDNFDSVMEAVDRVRQVVGPLLSQNLQFT 64
Db 3 ARNLVSGGAGYIGSHTVLQLLGGYSVVVDNLDNSAVSLQVRVKLAAB-HGERLSFH 61
QY 65 QGDLNRDDLEKLFKTTTFDAVIHPAGLKAVAESVAKPRRYPDFNLVGTINLYEFMAKN 124
Db 62 QVDLDRSALEKIFSETKFDVAIHPAGLKAVGESVEKELFYNNLNLTITLLEYAQRG 121
QY 125 CKKVFSSSATVYGOPKIPCEEDFKLOAMPYGTGKLFLEBIARDIOKABPEWKIILLR 184
Db 122 CKNLVFSSSATVYGSKPEVPCTEEPISALAPYGTGKLFIEICRDVYGSPEWKIILLR 181
QY 185 YFNPGAHESGKLGDPKGI PNMLPYYIOQVAVGSLTELNVYGHDPYTRDGSARDYIHV 244
```

Db 182 YFNFVGAHPSGDIGEDPRGIPNNLMFVQOVAVGRPHLTVFNDYNTKDGTVGRDYIHV 241
Qy 245 MDLADGHIAALRKLFETTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRRP 304
Db 242 IDLADGHIAALRKLEDCX-IGCEVNLGTGNGTSVLEMDAFKASGKKIPLVIAGRRP 300
Qy 305 DATEVYASTERAEKELGWKANGVVEEMCRDOWNAKNNPWGY 346
Db 301 DAEVYASTERAESELNWKAKYGIEMCRDLNWNASNNPYG 342

RESULT 12

AAG53259
ID AAG53259 standard; protein; 348 AA.

XX AC AAG53259;

XX DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 67789.

DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 31-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0130891P.

XX 30-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 04-MAY-1999; 99US-0132407P.

XX 05-MAY-1999; 99US-0132484P.

XX 05-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132863P.

XX 14-MAY-1999; 99US-0134256P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134370P.

XX 18-MAY-1999; 99US-0134768P.

XX 19-MAY-1999; 99US-0134941P.

XX 20-MAY-1999; 99US-0135124P.

XX 21-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.

XX 25-MAY-1999; 99US-0136021P.

XX 27-MAY-1999; 99US-0136392P.

XX 28-MAY-1999; 99US-0136782P.

XX 03-JUN-1999; 99US-0137222P.

XX 03-JUN-1999; 99US-0137528P.

XX 04-JUN-1999; 99US-0137502P.

XX 07-JUN-1999; 99US-0137724P.

XX 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 21-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 21-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.

PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135253P.
PR 24-MAY-1999; 99US-0135622P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148199P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.


```
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147119P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159323P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.

PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.1%; Score 1220; DB 3; Length 348;
Best Local Similarity 65.4%; Pred. No. 9e-108;
Matches 225; Conservative 51; Mismatches 66; Indels 2; Gaps 2;

QY 7 HILVTGGAGFTGTHVQLLKAGFSVSIIDNFNSVMEAVDRVQVVGVPLLSQNLQFTQG 66
DB 4 NILVTGGAGYIGSHVTLQLLLGGYTVVIDNLDNSSLVSIGRVKDLAD-HGQNLTVHQV 62
QY 67 DLNRDDLEKLFSTFDVAIHFAGLKAVASVAKPRRYFDNVLGVIINLYEFMAKYNCK 126
DB 63 DLDRKPALEKVFSETKFDVNHFAGLKAVGSVAKPLLYNNNLIAITLLEVMHAHGCK 122
QY 127 KMVFSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLLEEIARDIQKAEPEWKILRLRYF 186
DB 123 KLVFSSATVYGNWKEVCTEESPLSGNSPYGRTKLFIEDICRDVQGDPEWRILMLRYF 182
QY 187 NPVGAHESKLGEDPKGIPNKLMPYIQQVAVGRUTELNVYGHDPTRDGSIRDIYHMD 246
DB 183 NPVGAHPSGRIGEDPCGTPNNLMPYQVQVVGRLPNLKIYGTDTYTKDGTGVRDIHYVD 242
QY 247 LADGHIAALRLFTTENIGCTAYNLGTGRGTSVLEMTAFPEKASKGKIPVKLCPRRPGDA 306
DB 243 LADGHIFALQKLDDE-IGCEVYNLGTGKGTIVLEWDVAFKASKGMKIPVKVGRPGDA 301
QY 307 TEVYASTERAKELGWKANYGVVEEMCRDQWNAKNPMGYAGKP 350
DB 302 ETVYASTERELNWKANFGLEEMCRDQWNAKNPMGYGSSP 345

RESULT 15
AAG21994
ID AAG21994 standard; protein; 372 AA.
XX
AC AAG21994;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24755.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
```

PR 05-MAR-1999; 99US-01231802P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0126788P.
PR 25-MAR-1999; 99US-0126788P.
PR 29-MAR-1999; 99US-0126284P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130691P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139432P.
PR 18-JUN-1999; 99US-0139434P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140333P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145278P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147418P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149262P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149802P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154773P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.

```

PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.1%; Score 1220; DB 3; Length 372;
Best Local Similarity 65.4%; Pred. No. 9.9e-108;
Matches 225; Conservative 51; Mismatches 66; Indels 2; Gaps 2;

Qy 7 HILVTGGAGFIGTHVQLLKAGFSVSIIDNEDNSVMEAVDRVQVVGPLLSQNLQFTQG 66
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 NILVTGGAGVIGSHTVLQLLGYNVTVIDNLSLSVQIRVKDLAGD-HGQNLTVHQV 86
Qy 67 DLNRDDLEKLFKTTFDVIVHIFAGLKAVASVAKPRRYFDNLTGNTINLYEFMAKYNCK 126
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 DLDRKPALEKVFSETKFDVAVHIFAGLKAVGESVAKPLLYNNNLIAITITLLEVMAHGCK 146
Qy 127 KMFSSSATVYGOEKPCEDEFKIQAMNPYGETKLFLEIARDIQKAEPEWKIILLRYF 186
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
147 KLVFSSSATVYGPKEVPCFEESFLSGMSPYGTKLFIDICNDVQRGDPFWRIIMLRYF 206
Qy 187 NPVGAHESGKLGEDPKGIPNNLMPIYQQVAVGRLTELVYGHDPYPRDGSAIRDYIHVMD 246
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
207 NPVGHPGSRIGEDPCCPTNNLMPIYQQVVGRLPNLKIYGTDTTKDGTGVVDYIHVD 266
Qy 247 LADGHIAALAKLTETNIGCTAYNLGTGRTSVLENWTAPEKASGKKIPVKLCPRRPGDA 306
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
267 LADGHIFALQKLDDE-IGCEVNLGTGKTIVLEWDAFEKASGMKIPLVKVGRRPGDA 325
Qy 307 TEVYASTERAEKELGWKANYGVEEMCRDQWNAKNPWGYAGKP 350
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 ETYVASTEKAEELNWKANYGIEEMCRDQWNAKNPWGYAGKP 369

```

Search completed: March 22, 2004, 21:53:48
 Job time : 78 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 21:52:35 ; Search time 23 Seconds
(without alignments)
785.613 Million cell updates/sec

Title: US-09-913-064A-14

Perfect score: 1845

Sequence: 1 MVSSSHILVTGGAGFIGTH.....MCRDOWNWKNPNWGYAGRP 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
 - 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
 - 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
 - 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
 - 5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
 - 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	977.5	53.0	348	3	US-09-113-536-2
2	977.5	53.0	348	4	US-09-624-183-2
3	977.5	53.0	348	5	PCT-US95-05785-2
4	945	51.2	376	4	US-09-540-236-2180
5	926.5	50.2	345	4	US-09-543-681A-4948
6	918	49.8	363	4	US-09-489-039A-13195
7	893.5	48.4	338	4	US-09-580-929-4
8	633	34.3	469	4	US-09-107-532A-5936
9	624	33.8	341	4	US-09-134-000C-5524
10	619	33.6	329	4	US-09-134-000C-6238
11	553.5	30.0	359	4	US-09-634-955B-21
12	449.5	24.4	341	4	US-09-328-352-7330
13	334	18.1	329	4	US-09-107-532A-7038
14	318	17.2	344	4	US-09-489-039A-8087
15	311	16.9	99	4	US-09-424-311-3
16	308.5	16.7	347	4	US-09-543-681A-4531
17	306	16.6	339	4	US-09-489-039A-8139
18	301.5	16.3	339	4	US-09-252-991A-26600
19	278	15.1	341	4	US-09-580-929-3
20	278	15.1	352	4	US-09-580-929-2
21	272	14.7	290	4	US-09-134-000C-4233
22	268.5	14.6	91	4	US-09-424-311-4
23	266.5	14.4	276	4	US-09-134-000C-3713
24	262	14.2	346	4	US-09-328-352-5625
25	244	13.2	370	4	US-09-543-681A-4353
26	218.5	11.8	329	3	US-09-036-987A-26
27	218.5	11.8	329	3	US-09-370-700-26

Sequence 26, Appli
Sequence 23322, A
Sequence 19041, A
Sequence 5, Appli
Sequence 8053, Ap
Sequence 4, Appli
Sequence 6233, Ap
Sequence 15, Appl
Sequence 17, Appl
Sequence 15, Appl
Sequence 14, Appl
Sequence 35, Appl
Sequence 11, Appl
Sequence 10678, A
Sequence 7414, Ap
Sequence 7549, Ap
Sequence 17956, A
Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-113-536-2

; Sequence 2, Application US/09113536

; Patent No. 6153739

; GENERAL INFORMATION:

; APPLICANT: JI, H., ET AL.

; TITLE OF INVENTION: HUMAN UDP GALACTOSE-4-EPIMERASE

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILLILLAN,

; ADDRESSEE: CECCHI, STEWART & OLSTEIN

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/113,536

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/462,966

; FILING DATE: June 5, 1995

; APPLICATION NUMBER: PCT/US95/05785

; FILING DATE: 11 MAY 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: FERRARO, GREGORY D.

; REGISTRATION NUMBER: 36,134

; REFERENCE/DOCKET NUMBER: 325800-430

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 348 AMINO ACIDS

; TYPE: AMINO ACID

; STRANDEDNESS:

; TOPOLOGY: LINEAR

; MOLECULE TYPE: PROTEIN

US-09-113-536-2

Query Match 53.0%; Score 977.5; DB 3; Length 348;
Best Local Similarity 53.9%; Pred. NO. 4.2e-91;
Matches 188; Conservative 58; Mismatches 90; Indels 13; Gaps 4;

QY 124 -----NCKMVFSSATVYGQPEKIPCE---EDFKLOANMPYGRTKLFLEBIARDIQ 172
 Db 119 SALDETAKGVKVFSTDEVYGDLESIPISAFTEDPYNPSSPYGASKASSELLVRAHY 178
 QY 173 KAPEWKIILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQVAVGRLTETELNYGHDYPT 232
 Db 179 RAY-GLPAILLRYFNVYGYQSGRIGEDNGPEKIPILIONALOKGEPLVYGDYPT 237
 QY 233 RDGSAIRDYHVMDLA-DGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASG 291
 Db 238 PDGTQVRDTHVDEHARANELLALTR--GRACKGSEVYVNIQGGNEYSNLEVEATEKLLG 295
 QY 292 -----KKIPVKLCPRPGDATEVYASTERAELGKWKANY-GVEEMCRDQWNA 339
 Db 296 ELAPEKPHVKAKEDPATFVDDRPDDARYAADASKIKRELGWKPEVTNLEEGSLATVWNY 355
 QY 340 KNN 342
 Db 356 LEN 358

RESULT 12

US-09-328-352-7330
 ; Sequence 7330, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 7330
 ; LENGTH: 341
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-7330

Query Match 24.4%; Score 449.5; DB 4; Length 341;
 Best Local Similarity 31.2%; Pred. No. 2.5e-37;
 Matches 105; Conservative 74; Mismatches 147; Indels 11; Gaps 5;
 QY 8 ILVTGAGIGTHTVOLLKAGFSVSIIDFNFSVMEAVDRVRQVVGPLLSONLOFTQGD 67
 Db 8 ILVTGGLGFIGSHIALSLMAQGVVLDNLANSTLQTLERLEYISG---MYVPFVKLD 63
 QY 68 LRNRDLEKLFSTTDFDAVIHFAGLKAVASVAKPRRYFDENLVGTINLYEFMAKYNCK 127
 Db 64 VENTPALNVFQYSDAIVIHAGFKSIEESNLKPELYNDVNSCINLLRAMQRTGVH 123
 QY 128 MVFSSATVYGQPEKIPCEEDFKLOANMPYGRTKLFLEBIARDIQKABPEWKIILLRYF 186
 Db 124 FHLSSLAVYKSGSLSQSETDFNAYPNFYTKSQQMIEIIRDTYKIDHEWKIAILRLS 183
 QY 187 NPVGASHGKLGEDPKGIPNNLMPYIQVAVR--LTELNVYGHDPYPTDGSARDYTHV 244
 Db 184 NIVGAFEHVGLGYAQLFKNLIVPLAMQVAAQCRDLIELQ---DOSSTEDHTVRSFLHV 240
 QY 245 MDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRPG 304
 Db 241 LDVCEAVSSTLHL--RDQSHCCAFNAIHEQVHSIRQLLDEISQVTAEPITQSAIYKHE 299
 QY 305 DATEVYASTERAELGKWKANYGVEEMCRDQWNAK 341
 Db 300 ELAQIGANTGKATLLQWPKRTLQWIEDERFYQN 336

RESULT 13

US-09-107-532A-7038
 ; Sequence 7038, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 7038:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 329 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...329
 SEQUENCE DESCRIPTION: SEQ ID NO: 7038:
 US-09-107-532A-7038
 Query Match 18.1%; Score 334; DB 4; Length 329;
 Best Local Similarity 27.1%; Pred. No. 1.4e-25;
 Matches 90; Conservative 72; Mismatches 124; Indels 46; Gaps 10;
 QY 9 LVTGGAGFIGTHTVOLLKAGFSVSIIDFNFSVMEAVDRVRQVVGPLLSONLOFTQGD 68
 Db 19 LITGGAGFIGS--TLANYLGKNNVVYVDDLSMGKKNLDE-----SKHITFIEGDV 68
 QY 69 RNRDLEKLFSTTDFDAVIHFAGLKAVASVAKPRRYFDENLVGTINLYEFMAKYNCK 126
 Db 69 ADKNLKEKVKWEHFYIYIPLAAVASVADSVVERPLETHRVNFPDALLLELVKYSGLK 128
 QY 127 KMFSSATVYGQPEKIPCEEDFKLOANMPYGRTKLFLEBIARDIQKABPEWKIILLRYF 186
 Db 129 RLVFSSAAVYGDPELTPKKEESVIRPLTPYADKFAAEQYVLDYCHLY-DVPTSAVRFF 187
 QY 187 NPVGASHGKLGEDPKGIPNNLMPYIQVAV-----GRLTENLVYGHDPYPTDGS 236
 Db 188 NVTGPQN-----PNS--PYGVSILVDYRKKQLAGEKTEFTLFG-----DGS 229
 QY 237 AIRDYIHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPV 296
 Db 230 QSRDFVYIEDV---IQAALLVAKBEKALQQQNVGTGKSTLLELIHSDIQLGTAL 285
 QY 297 KLCFRPGDATEVYASTERAELGKWKANYGV 328
 Db 286 KYAERSGDIRDSLADISKI-RSLGYQPKFDI 316

RESULT 14
US-09-489-039A-8087
; Sequence 8087, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8087
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8087

Query Match 17.2%; Score 318; DB 4; Length 344;
Best Local Similarity 28.7%; Pred. No. 6.4e-24;
Matches 102; Conservative 58; Mismatches 144; Indels 52; Gaps 12;
Qy 9 LVTGAGFIGTHVTVQLLKAGFSVSIIDN----FDRSVMEA-VDRVQVVGPLLSONLQF 63
Db 14 LVTGAGFIGTHVTVQLLKAGFSVSIIDN----FDRSVMEA-VDRVQVVGPLLSONLQF 63
Qy 64 TQGLNRDDLEKLFSTTDDAVIHFAGLKAVAEVAKPRRYFDNVLVGTINLYEFMAKY 123
Db 67 QQLDLADREGMAKLFATEQDFRVHILAAQAGVRYSLNPYADANLGMVNLILEGCRHT 126
Qy 124 NCKWVFSSATVYQGEKIP-CEEDFKLQAMPYGRKLFLEIARDIQAEPEWKIL 182
Db 127 KVHLVYASSSVYGLNRKPFSTEDSVDFHPSVLYAATKKANLMAHTYSHL---YGIPT 183
Qy 183 --LRYFNPVGAHSGKLGEDPKGIPN-NLMPYIQQAVAGLTELNVYGHDPYTRDGSAIR 239
Db 184 TGLRFFTVYG-----PWGRPDNALPKFYKAMLEK--SIDVNY-----GVMKR 225
Qy 240 DYHVMADGHIHAAKLFETENIGCT-----AYNLGTGRTSVLEWVTA 285
Db 226 DFTYIDIVEA-VVRVQDVIPQANADWTVESGSPATSSAPRYVYVNIIGNSSPVELMDYTA 284
Qy 286 FEXASGKKIPVKLCPRPGDATEVYASTERAERKELGWKANYGVEMCRDQWNAKN 341
Db 285 LEEALGMEAKKNMPIOQGDVLDTSADTQALYDLVGFQPTSVKGVNFWFYKD 340

RESULT 15
US-09-424-311-3
; Sequence 3, Application US/09424311
; Patent No. 6372477
; GENERAL INFORMATION:
; APPLICANT: JORSSBOE, Morten
; BRUNSTEDT, Janne
; PETERSEN, Steen Guldager
; TITLE OF INVENTION: CLONING OF UDP-GALACTOSE EPIMERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSES: FOLEY & LARNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/424,311
FILING DATE: 24-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/IB98/00886
FILING DATE: 27-MAY-1998
APPLICATION NUMBER: GB 9710991.2
FILING DATE: 28-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 078883/0114
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-424-311-3

Query Match 16.9%; Score 311; DB 4; Length 99;
Best Local Similarity 70.2%; Pred. No. 4.8e-24;
Matches 66; Conservative 8; Mismatches 20; Indels 0; Gaps 0;
Qy 2 VSSSHILVTGAGFIGTHVTVQLLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSONL 61
Db 6 MASGETILVTGAGFIGSHVTVQLLKQGFHVSIIIDNLYNSVIDAVHVRVLLVGPILLSNL 65
Qy 62 QFTQGLNRDDLEKLFSTTDDAVIHFAGLKAV 95
Db 66 HPHGDLNRHDDLLFSTKTFDAVIQLAGPKGV 99

Search completed: March 22, 2004, 21:56:31
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 21:55:36 ; Search time 63 Seconds
(without alignments)
1438.641 Million cell updates/sec

Title: US-09-913-064A-14
Perfect score: 1845
Sequence: 1 MVSSSQHLVTGGAGFIGTH.....MCRQWNAKNPNWYACKP 350

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10E_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10F_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10G_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1840	99.7	350	12	US-10-424-599-179426
2	1762	95.5	363	12	US-10-425-114-47644
3	1486	80.5	351	14	US-10-060-275-2
4	1198.5	65.0	334	12	US-10-425-114-46573
5	1196.5	64.9	334	12	US-10-425-114-39413
6	1192	64.6	391	12	US-10-424-599-232009
7	1192	64.6	408	12	US-10-425-114-55605
8	1190	64.5	408	12	US-10-425-114-56199
9	1183.5	64.1	441	12	US-10-424-599-232008
10	1148	62.2	403	12	US-10-425-114-59200
11	1141	61.8	382	12	US-10-425-114-61106
12	1071.5	58.1	333	12	US-10-425-114-58229
13	1040	56.4	326	12	US-10-425-114-41119
14	986.5	53.5	348	9	US-09-318-271-10
15	986.5	53.5	348	14	US-10-303-664A-18

16	977.5	53.0	348	13	US-10-201-298-2	Sequence 2, Appli
17	918.5	49.8	350	13	US-10-108-605-65	Sequence 65, Appl
18	893.5	48.4	338	9	US-09-318-271-8	Sequence 8, Appli
19	887.5	48.1	336	15	US-10-429-812-10	Sequence 10, Appl
20	878.5	47.6	699	15	US-10-369-493-1430	Sequence 1430, Ap
21	836.5	45.3	255	12	US-10-424-599-148894	Sequence 148894,
22	775.5	42.0	207	12	US-10-424-599-232005	Sequence 232005,
23	775.5	42.0	207	12	US-10-425-114-45766	Sequence 45766, A
24	761	41.2	238	12	US-10-424-599-232010	Sequence 232010
25	751	40.7	286	12	US-10-425-114-55233	Sequence 55233, A
26	683.5	37.0	216	12	US-10-425-114-44206	Sequence 44206, A
27	645.5	35.0	170	12	US-10-425-114-43920	Sequence 43920, A
28	634	34.4	415	12	US-10-424-599-273943	Sequence 273943,
29	625.5	33.9	513	12	US-10-425-114-62570	Sequence 62570, A
30	616.5	33.4	497	12	US-10-425-114-67924	Sequence 67924, A
31	615	33.3	327	15	US-10-369-493-14762	Sequence 14762, A
32	614	33.3	209	12	US-10-425-114-47486	Sequence 47486, A
33	612.5	33.2	438	12	US-10-425-114-60686	Sequence 60686, A
34	608	33.0	327	15	US-10-369-493-16982	Sequence 16982, A
35	569	30.8	325	15	US-10-369-493-19845	Sequence 19845, A
36	562.5	30.5	319	14	US-10-156-761-11112	Sequence 11112, A
37	557	30.2	423	12	US-10-424-599-273945	Sequence 273945,
38	553.5	30.0	359	14	US-10-172-585-21	Sequence 21, Appl
39	544.5	29.5	329	9	US-09-738-626-5610	Sequence 5610, Ap
40	534.5	29.0	325	15	US-10-369-493-8085	Sequence 8085, Ap
41	530.5	28.8	326	14	US-10-156-761-9201	Sequence 9201, Ap
42	526.5	28.5	321	15	US-10-369-493-10165	Sequence 10165, A
43	501.5	27.2	326	14	US-10-156-761-12620	Sequence 12620, A
44	487.5	27.0	258	12	US-10-425-114-63671	Sequence 63671, A
45	491.5	26.6	317	15	US-10-369-493-21857	Sequence 21857, A

ALIGNMENTS

RESULT 1
US-10-424-599-179426
; Sequence 179426, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 179426
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_133037C.1.pep
US-10-424-599-179426

Query Match	99.7%	Score	1840	DB	12	Length	350
Best Local Similarity	99.7%	Pred. No.	8.2e-187				
Matches	349	Conservative	0	Mismatches	1	Indels	0
						Gaps	0
Qy	1	MVSSSQHLVTGGAGFIGTHVQVLLKAGFSVSIIDFNSVMEAVDRVQVVGFLLSQN	60				
Db	1	MVSSSQHLVTGGAGFIGTHVQVLLKAGFSVSIIDFNSVMEAVDRVQVVGFLLSQN	60				
Qy	61	LQFTQGLNRDLEKLFSTKTFDAVTHFAGLKAVASVAKPRRYFDFNLVTGTLNLEFYM	120				
Db	61	LQFTQGLNRDLEKLFSTKTFDAVTHFAGLKAVASVAKPRRYFDFNLVTGTLNLEFYM	120				
Qy	121	AKYNCKKVVSSSATVVGQPEKIPCEEDFKLQAMNPYGRYTKLFLEEIARDIQKAEPSWKI	180				
Db	121	AKYNCKKVVSSSATVVGQPEKIPCEEDFKLQAMNPYGRYTKLFLEEIARDIQKAEPSWKI	180				

QY 181 ILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTFLNLYGHDYPTDGSARD 240
Db 181 ILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTFLNLYGHDYPTDGSARD 240
QY 241 YHVMDLADGHI AALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 300
Db 241 YHVMDLADGHI AALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 300
QY 301 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNAKKNPWGYAGKP 350
Db 301 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNAKKNPWGYAGKP 350

RESULT 2

US-10-425-114-47644
; Sequence 47644, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47644
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700787631_FLI.pbp
US-10-425-114-47644

Query Match 95.5%; Score 1762; DB 12; Length 363;
Best Local Similarity 95.1%; Pred. No. 1.7e-178;
Matches 333; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 MYSSOHLVTGGAGFIGTHTVQVLLKAGFSVSIIDNFNSVMEAVDRVRQVVGPLLSQN 60
Db 14 WTSSQKILVTGGAGFIGTHTVQVLLKAGFSVSIIDNFNSVMEAVDRVRQVVGPLLSQN 73
QY 61 LOFTQGLRNRDDEKLFSTTTPDAVTHPAGLKAVAESVAKPRRYDFDNLVGTINLYEFM 120
Db 74 LOFTQGLRNRDDEKLFSTTTPDAVTHPAGLKAVAESVAKPRRYDFDNLVGTINLYEFM 133
QY 121 AKYCKKRVFSSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQKAPENKI 180
Db 134 AKYCKKRVFSSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQKAPENKI 193
QY 181 ILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTFLNLYGHDYPTDGSARD 240
Db 194 ILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTFLNLYGHDYPTDGSARD 253
QY 241 YHVMDLADGHI AALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 300
Db 254 YHVMDLADGHI AALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 313
QY 301 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNAKKNPWGYAGKP 350
Db 314 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNAKKNPWGYAGKP 363

RESULT 3

US-10-060-275-2
; Sequence 2, Application US/10060275
; Publication No. US20030073828A1
; GENERAL INFORMATION:

; APPLICANT: DAI, ZIYU
; APPLICANT: SHI, LIFANG
; APPLICANT: HOOKER, BRIAN S.
; TITLE OF INVENTION: EPIMERASE GENE AND USE THEREOF
; FILE REFERENCE: 059440-0143
; CURRENT APPLICATION NUMBER: US/10/060,275
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/265,311
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-10-060-275-2

Query Match 80.5%; Score 1486; DB 14; Length 351;
Best Local Similarity 78.2%; Pred. No. 4e-149;
Matches 273; Conservative 33; Mismatches 43; Indels 0; Gaps 0;

QY 2 VSSSOHLVTGGAGFIGTHTVQVLLKAGFSVSIIDNFNSVMEAVDRVRQVVGPLLSQN 61
Db 3 VQCOENILVTGGAGFIGTHTVQVLLNEGPKVTIIDNFHNSVEEAVDRVRELVGPLSQNL 62
QY 62 QFTQGLRNRDDEKLFSTTTPDAVTHPAGLKAVAESVAKPRRYDFDNLVGTINLYEFM 121
Db 63 EFLGDIRNRDDEKLFSTTTPDAVTHPAGLKAVGESVVOPLYFENNIGSITILSVMA 122
QY 122 KYNCKKRVFSSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQKAPENKI 191
Db 123 KYNCKKRVFSSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQKAPENKI 192
QY 182 LLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTFLNLYGHDYPTDGSARD 241
Db 183 LLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTFLNLYGHDYPTDGSARD 242
QY 242 IHVMDLADGHI AALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 301
Db 243 IHVMDLADGHI AALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 302
QY 302 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNAKKNPWGYAGKP 350
Db 303 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNAKKNPWGYAGKP 351

RESULT 4

US-10-425-114-46573
; Sequence 46573, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46573
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700445105_FLI.pbp
US-10-425-114-46573

Query Match 65.0%; Score 1198.5; DB 12; Length 394;
Best Local Similarity 66.0%; Pred. No. 1.9e-118;

```
Matches 231; Conservative 44; Mismatches 72; Indels 3; Gaps 3;
QY 1 MVSS-SQHILVTGGAGTCTHTVQVOLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQ 59
Db 40 MVSALVLTILVTGGAGTCTHTVQVOLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQ 98
QY 60 NLOFTQGLNRDLEKLFSTTFDAVTHFAGLKAVASVAKPRRYFDNLTGTLNLYEF 119
Db 99 NLVFKVLDLRDRHALVDIFSSHREFAVTHFAGLKAVGESVHKPLLYDNNLVGTLTLLV 158
QY 120 MAKYNCKMVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLEETARDIOKAEPEWK 179
Db 159 MAANGCKLVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLEETARDIOKAEPEWK 218
QY 180 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQVAVAGRLTFLNLYGHDYPTDGSAR 239
Db 219 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQVAVAGRLTFLNLYGHDYPTDGSAR 278
QY 240 DYHWMDLADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLVF 298
Db 279 DYHWMDLADGHIAALRKLYEDSKTGCEVYNLTGKTSVLEMTAFKASGKKIPLVF 338
QY 299 CPERPGDAEVYASTRAEKELGWKANYGVEMCRDQWNAKNNPWGYAG 348
Db 339 AGRPGDAEIVYATAKAEKELGWKANYGVEMCRDQWNAKNNPWGYAG 388

RESULT 5
US-10-425-114-39413
; Sequence 39413, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 39413
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700103191_FLI.pep
US-10-425-114-39413
Query Match 64.9%; Score 1196.5; DB 12; Length 394;
Best Local Similarity 65.7%; Pred. No. 3.1e-118;
Matches 230; Conservative 45; Mismatches 72; Indels 3; Gaps 3;
QY 1 MVSS-SQHILVTGGAGTCTHTVQVOLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQ 59
Db 40 MVSALVLTILVTGGAGTCTHTVQVOLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQ 98
QY 60 NLOFTQGLNRDLEKLFSTTFDAVTHFAGLKAVASVAKPRRYFDNLTGTLNLYEF 119
Db 99 NLVFKVLDLRDRHALVDIFSSHREFAVTHFAGLKAVGESVHKPLLYDNNLVGTLTLLV 158
QY 120 MAKYNCKMVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLEETARDIOKAEPEWK 179
Db 159 MAANGCKLVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLEETARDIOKAEPEWK 218
QY 180 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQVAVAGRLTFLNLYGHDYPTDGSAR 239
Db 219 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQVAVAGRLTFLNLYGHDYPTDGSAR 278
QY 240 DYHWMDLADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLVF 298
```

```
Db 279 DYHWMDLADGHIAALRKLYEDSKTGCEVYNLTGKTSVLEMTAFKASGKKIPLVF 338
QY 299 CPERPGDAEVYASTRAEKELGWKANYGVEMCRDQWNAKNNPWGYAG 348
Db 339 AGRPGDAEIVYATAKAEKELGWKANYGVEMCRDQWNAKNNPWGYAG 388

RESULT 6
US-10-424-599-232009
; Sequence 232009, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285694
; SEQ ID NO 232009
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51526C.1.pep
US-10-424-599-232009
Query Match 64.6%; Score 1192; DB 12; Length 391;
Best Local Similarity 66.4%; Pred. No. 9.2e-118;
Matches 225; Conservative
QY 8 ILVTGGAGTCTHTVQVOLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQNLQFTQGD 67
Db 48 VLVTGGAGTCTHTVQVOLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQNLQFTQGD 106
QY 68 LRNRDDLEKLFSTTFDAVTHFAGLKAVASVAKPRRYFDNLTGTLNLYEFMAKYNCK 127
Db 107 LRDRALDQIFSSQTFDAVTHFAGLKAVGESVQKPLLYNNLTGTLTLLVMAAGCK 166
QY 128 MVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLEETARDIOKAEPEWKIILLYEFN 187
Db 167 LVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLEETARDIOKAEPEWKIILLYEFN 226
QY 188 PVGAHSGKLGEDPKGIPNNLMPYIQVAVAGRLTFLNLYGHDYPTDGSARIDYHWM 247
Db 227 PVGAHSGKLGEDPKGIPNNLMPYIQVAVAGRLTFLNLYGHDYPTDGSARIDYHWM 286
QY 248 ADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLVLCRRPGDAT 307
Db 287 ADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLVLCRRPGDAT 346
QY 308 EYASTRAEKELGWKANYGVEMCRDQWNAKNNPWGY 346
Db 346 EYASTRAEKELGWKANYGVEMCRDQWNAKNNPWGY 384

RESULT 7
US-10-425-114-55605
; Sequence 55605, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55605
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700103191_FLI.pep
US-10-425-114-55605
Query Match 64.9%; Score 1196.5; DB 12; Length 394;
Best Local Similarity 65.7%; Pred. No. 3.1e-118;
Matches 230; Conservative 45; Mismatches 72; Indels 3; Gaps 3;
QY 1 MVSS-SQHILVTGGAGTCTHTVQVOLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQ 59
Db 40 MVSALVLTILVTGGAGTCTHTVQVOLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQ 98
QY 60 NLOFTQGLNRDLEKLFSTTFDAVTHFAGLKAVASVAKPRRYFDNLTGTLNLYEF 119
Db 99 NLVFKVLDLRDRHALVDIFSSHREFAVTHFAGLKAVGESVHKPLLYDNNLVGTLTLLV 158
QY 120 MAKYNCKMVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLEETARDIOKAEPEWK 179
Db 159 MAANGCKLVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLEETARDIOKAEPEWK 218
QY 180 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQVAVAGRLTFLNLYGHDYPTDGSAR 239
Db 219 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQVAVAGRLTFLNLYGHDYPTDGSAR 278
QY 240 DYHWMDLADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLVF 298
```

FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 55605
LENGTH: 408
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: UC-GMROINOI018B10_FLI.pap
US-10-425-114-55605

Query Match 64.6%; Score 1192; DB 12; Length 408;
Best Local Similarity 66.4%; Pred. No. 9.8e-118;
Matches 225; Conservative 44; Mismatches 68; Indels 2; Gaps 2;

QY 8 ILVTGGAGFIGTHTVQLLKAGFSVSIIDNPDNSVMEAVDRVQVGPVLLSONLQFTQGD 67
DB 65 VLVTGGAGVIGSHTVQLLGGFRAVVDLNLNSSEVAIHRVRELAGE-FGNLSPHKVD 123
QY 68 LRNRDLEKLFSTQTFDAVIHFAGLKAVASVAKPRRYDFNLVGTINLYEFMAKYNCK 127
DB 124 LRDRALDQIFSTQTFDAVIHFAGLKAVGESVOKPLLYNNNLGTITLLEVMAAHGCK 183
QY 128 MYPSSATVYGOPEKIPCEEDFKLOAMPYGRTKLFLEIARDIOKAEPEWKIILLRYFN 187
DB 184 LVFSSATVYGPKEVPEETEEFPLSAMNPGYGRTKLIIIEICRDVCAEPDCKIILLRYFN 243
QY 188 PVGAHSGKLGEDPDKGIPNNLMPYIQQAVAGRLTELNVYGHDPYTRDGSARDYIHVMDL 247
DB 244 PVGAHPSGYIGEDPRGIPNNLMPFVQVAVGRRPALTVFGNDYNTSDGTGVRDYIHVVDL 303
QY 248 ADGHTAALRKLFETENIGCTAYNLGTGRGTSVLEWYTAPEKASGKKIPVKLCPRRPGDAT 307
DB 304 ADGHTAALRLK-DEPNIGCEVNLGTGKTSVLEWYRAPEFMAAGKIPLVMAGRRPDGA 362
QY 308 EYVASTERAEKELGWKANYGVEMCRDQWNAKNPNWGY 346
DB 363 IVYASTKKAERELKWKAKYGIDEMCRDQWNAKNPNWGY 401

RESULT 8
US-10-425-114-56199
Sequence 56199, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 56199
LENGTH: 408
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: UC-GMROFIC090E09_FLI.pap
US-10-425-114-56199

Query Match 64.5%; Score 1190; DB 12; Length 408;
Best Local Similarity 66.4%; Pred. No. 1.6e-117;
Matches 225; Conservative 44; Mismatches 68; Indels 2; Gaps 2;

QY 8 ILVTGGAGFIGTHTVQLLKAGFSVSIIDNPDNSVMEAVDRVQVGPVLLSONLQFTQGD 67
DB 65 VLVTGGAGVIGSHTVQLLGGFRAVVDLNLNSSEVAIHRVRELAGE-FGNLSPHKVD 123

QY 68 LRNRDLEKLFSTQTFDAVIHFAGLKAVASVAKPRRYDFNLVGTINLYEFMAKYNCK 127
DB 124 LRDRALDQIFSTQTFDAVIHFAGLKAVGESVOKPLLYNNNLGTITLLEVMAAHGCK 183
QY 128 MYPSSATVYGOPEKIPCEEDFKLOAMPYGRTKLFLEIARDIOKAEPEWKIILLRYFN 187
DB 184 LVFSSATVYGPKEVPEETEEFPLSAMNPGYGRTKLIIIEICRDVCAEPDCKIILLRYFN 243
QY 188 PVGAHSGKLGEDPDKGIPNNLMPYIQQAVAGRLTELNVYGHDPYTRDGSARDYIHVMDL 247
DB 244 PVGAHPSGYIGEDPRGIPNNLMPFVQVAVGRRPALTVFGNDYNTSDGTGVRDYIHVVDL 303
QY 248 ADGHTAALRKLFETENIGCTAYNLGTGRGTSVLEWYTAPEKASGKKIPVKLCPRRPGDAT 307
DB 304 ADGHTAALRLK-DEPNIGCEVNLGTGKTSVLEWYRAPEFMAAGKIPLVMAGRRPDGA 362
QY 308 EYVASTERAEKELGWKANYGVEMCRDQWNAKNPNWGY 346
DB 363 IVYASTKKAERELKWKAKYGIDEMCRDQWNAKNPNWGY 401

RESULT 9
US-10-424-599-232008
Sequence 232008, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 232008
LENGTH: 441
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_51525C.1.pap
US-10-424-599-232008

Query Match 64.1%; Score 1183.5; DB 12; Length 441;
Best Local Similarity 53.6%; Pred. No. 8.9e-117;
Matches 232; Conservative 47; Mismatches 61; Indels 93; Gaps 4;

QY 5 SCHLVTGGAGFIGTHTVQLLKAGFSVSIIDNPDNSVMEAVDRVQVGPVLLSONLQFT 64
DB 3 SQSLVTGGAGYIGSHTVQLLGGFRAVVDNPNDSSETAINRVRELAGE-FANLUSFS 61
QY 65 QGDLNRDLEKLFSTQTFDAVIHFAGLKAVASVAKPRRYDFNLVGTINLYEFMAK-- 122
DB 62 KLDLRDRAALEKIFSTNKFDPAVIHFAGLKAVGSVDKPLLYFDNNLIGTIVLFEVMAAHG 121
QY 123 ----- 122
DB 122 CKLVPSSATVYGPKEVPEETEEFPLSATNPYGRTKLIIIEICRDIYRADSDWKVILLR 181
QY 123 -----YN-----CKKVPSSATVYGOPEKIPCEEDFKLOA 153
DB 182 YNESVQOPLLYNNNLGTITLLEVMAAHGCKKLVFSSATVYGPKEVPEETEEFPLSA 241
QY 154 MNPYGRTKLFLEIARDIOKAEPEWKIILLRYNPNVGAHSGKLGEDPKGIPNNLMPYIQ 213
DB 242 MNPYGRTKLIIIEICRDVCAEPDCKIILLRYNPNVGAHPSGCGIGEDPRGIPNNLMPFVQ 301
QY 214 QVAVGRLTELNVYGHDPYTRDGSARDYIHVMDLADGHTAALRKLFETENIGCTAYNLGT 273
DB 302 QVAVGRRPALTVFGNDYNTSDGTGVRDYIHVMDLADGHTAALRLK-LDDPNIGCEVNLGT 360

QY 274 GRTSVLEWYATEKASGKKIPVKLCPRPGDTEYVYATERAEKELGWKANYGVEMCR 333
DB 361 GKGTSVLEWYATEKASGKKIPVKLCPRPGDTEYVYATERAEKELGWKANYGVEMCR 420
QY 334 DOWNWAKNPGY 346
DB 421 DOWNWAKNPGY 433

RESULT 10
US-10-425-114-59200
; Sequence 59200, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59200
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700260760_FLI.pep
US-10-425-114-59200

Query Match 62.2%; Score 1148; DB 12; Length 403;
Best Local Similarity 62.4%; Pred. No. 4.7e-113;
Matches 217; Conservative 44; Mismatches 63; Indels 24; Gaps 2;

QY 3 SSSOHILVTGGAGTGTTHVQVOLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSQLQ 62
DB 13 SAGTSLVTGGAGTGTTHVQVOLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSQLQ 49
QY 63 FTQGLRNDDLEKLFSTFTDAVTHFAGLKAVASVAKPRRYDFNLVGTINLYEFMAK 122
DB 50 FHKIDLRDKGALEWVFASTFEAVTHFAGLKAVASVAKPRRYDFNLVGTINLYEFMAK 109
QY 123 YNCKMVFSSATVYGOPEKIPCEDEPKLQAMPYGRGTLFLEEIARDIOKAEPEWKIIL 182
DB 110 HGCKLVFSSAAVYGSPPKNSPCTEDFPLTPNPNYQVAVGRRPALTVLGNDAVTRDGTGVRDYI 169
QY 183 LRYENPVGAHSGKLGEDPKGIPNNLMPYIQQVAVGRRLTELNVYGHDPTRDGSARDYI 242
DB 170 LRYENPVGAHSGKLGEDPKGIPNNLMPYIQQVAVGRRPALTVLGNDAVTRDGTGVRDYI 229
QY 243 HVMDLADGHTAALRKLFTTENIGCTAYNLGRTGTVLEMTATFEKASGKKIPVKLCPRR 302
DB 230 HVMDLADGHTAALRKLFTTENIGCTAYNLGRTGTVLEMTATFEKASGKKIPVKLCPRR 289
QY 303 PGDATEVYATERAEKELGWKANYGVEMCRDQWNAKNNPWGYAGKP 350
DB 290 PGDATEVYATERAEKELGWKANYGVEMCRDQWNAKNNPWGYAGKP 336

RESULT 11
US-10-425-114-61106
; Sequence 61106, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61106
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3332-019-F4_FLI.pep
US-10-425-114-61106

Query Match 61.8%; Score 1141; DB 12; Length 382;
Best Local Similarity 62.0%; Pred. No. 2.4e-112;
Matches 214; Conservative 54; Mismatches 75; Indels 2; Gaps 2;

QY 4 SSOHILVTGGAGTGTTHVQVOLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSQLQ 63
DB 27 SSRVLVTGGAGTGTTHVQVOLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSQLQ 85
QY 64 TQGLRNDDLEKLFSTFTDAVTHFAGLKAVASVAKPRRYDFNLVGTINLYEFMAK 123
DB 86 HKVDIRDKXGLEKVFSTFEAVTHFAGLKAVASVAKPRRYDFNLVGTINLYEFMAK 145
QY 124 NCKMVFSSATVYGOPEKIPCEDEPKLQAMPYGRGTLFLEEIARDIOKAEPEWKIIL 183
DB 146 GKCKLVFSSAAVYGSPPKNSPCTEDFPLTPNPNYQVAVGRRPALTVLGNDAVTRDGTGVRDYI 205
QY 184 RYENPVGAHSGKLGEDPKGIPNNLMPYIQQVAVGRRLTELNVYGHDPTRDGSARDYI 243
DB 206 RYENPVGAHSGKLGEDPKGIPNNLMPYIQQVAVGRRPALTVLGNDAVTRDGTGVRDYI 265
QY 244 VMDLADGHTAALRKLFTTENIGCTAYNLGRTGTVLEMTATFEKASGKKIPVKLCPRR 302
DB 266 VMDLADGHTAALRKLFTTENIGCTAYNLGRTGTVLEMTATFEKASGKKIPVKLCPRR 325
QY 303 PGDATEVYATERAEKELGWKANYGVEMCRDQWNAKNNPWGYAGKP 347
DB 326 PGDATEVYATERAEKELGWKANYGVEMCRDQWNAKNNPWGYAGKP 370

RESULT 12
US-10-425-114-58229
; Sequence 58229, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58229
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3592-035-C3_FLI.pep
US-10-425-114-58229

Query Match 58.1%; Score 1071.5; DB 12; Length 333;
Best Local Similarity 64.3%; Pred. No. 4.8e-105;
Matches 207; Conservative 45; Mismatches 69; Indels 1; Gaps 1;

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 21:44:15 ; Search time 27 Seconds
(without alignments)
1246.927 Million cell updates/sec

Title: US-09-913-064A-14
Perfect score: 1845
Sequence: 1 MYSSSQHILVTGGAGFIGTH.....MCRDQWNAKNPWGAGKP 350

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1671	90.6	350	2 T06526	UDP-glucose 4-epime
2	1578	85.5	351	2 B86261	UDP-glucose 4-epime
3	1571	85.1	351	2 S62783	UDP-glucose 4-epime
4	1522	82.5	353	2 D96657	hypothetical prote
5	1500	81.3	354	2 T10496	UDP-glucose 4-epime
6	1246	67.5	350	2 T10498	UDP-glucose 4-epime
7	1229	66.6	350	2 T01881	UDP-glucose 4-epime
8	1229	66.6	351	2 T04291	probable UDP-glucos
9	1208	65.3	339	2 T08911	UDP-glucose 4-epime
10	1038	56.3	338	2 D59628	UDP-glucose 4-epime
11	983	53.3	338	2 F82419	UDP-glucose 4-epim
12	971.5	52.7	335	2 AC3357	UDP-glucose 4-epime
13	966	52.4	339	2 B35187	UDP-glucose 4-epim
14	964	52.2	339	2 C98034	UDP-glucose 4-epime
15	962	52.1	347	2 T19989	hypothetical prote
16	943	51.1	347	2 S11223	UDP-glucose 4-epime
17	941	51.0	336	2 D95213	UDP-glucose 4-epim
18	928.5	50.3	339	2 F82014	UDP-glucose 4-epime
19	920.5	49.9	338	1 A64063	UDP-glucose 4-epime
20	916.5	49.7	339	1 S39638	UDP-glucose 4-epime
21	916.5	49.7	339	2 S42430	UDP-glucose 4-epime
22	910.5	49.3	337	1 A36931	UDP-glucose 4-epime
23	910.5	49.3	338	2 AF0594	UDP-glucose 4-epime
24	908	49.2	326	2 E98077	UDP-glucose 4-epime
25	904.5	49.0	338	1 S51328	UDP-glucose 4-epime
26	904.5	49.0	338	1 S34984	UDP-glucose 4-epime
27	894.5	48.5	336	1 S70744	UDP-glucose 4-epime
28	894.5	48.5	338	2 C90727	UDP-galactose-4-ep
29	894.5	48.5	338	2 D85578	UDP-galactose-4-ep

ALIGNMENTS

RESULT 1

T06526
UDP-glucose 4-epimerase (EC 5.1.3.2) - garden pea
N/Alternate names: UDP-galactose 4-epimerase
C/Species: Pisum sativum (Garden pea)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jan-2000
C/Accession: T06526
R/Slocum, R.D.; Lake, M.R.; Williamson, C.L.
Submitted to the EMBL Data Library, July 1995
A/Description: Cloning and characterization of a UDP-galactose-4-epimerase ('galactowal')
A/Reference number: Z15735
A/Accession: T06526
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-350 <SLO>
A/Cross-references: EMBL:U31544; NID:g1173554; PIDN:AAA86532.1; PID:g1173555
C/Genetics:
C/Gene: gale
C/Superfamily: Escherichia coli UDP-glucose 4-epimerase; UDP-glucose 4-epimerase homology
C/Keywords: galactose metabolism; isomerase
F/8-346/Domain: UDP-glucose 4-epimerase homology <UDP>

Query Match 90.6%; Score 1671; DB 2; Length 350;

Best Local Similarity 90.0%; Pred. No. 2e-129;
Matches 315; Conservative 16; Mismatches 19; Indels 0; Gaps 0;

QY	1	MYSSSQHILVTGGAGFIGTHVTVOLLKAGFSVSIIDNFNSVMEAVDRYRVGVGPLLSQN	60
Db	1	MVASSQKILVTGSAGFIGTHVTVOLLNNGFNVSIIIDNFNSVMEAVRVRVGVGSLNSQN	60
QY	61	LQFTQGLNRDDLEKLFSTTFDAVIHFAGLKAVESVAKPRRYDFNLVGTINLYEPM	120
Db	61	LEFTGLDRNKDDLEKLFSTKSDAVIHFAGLKAVGESVENPRRYFDNLLVGTINLYEVM	120
QY	121	AKYNCKWVSSSTVYGOPEKIPCEDEKLOANNYGTGKLFLEIARDIOKAEPEWKI	180
Db	121	AKNCKWVSSSTATVYGOPEKIPCEDEKLOANNYGTGKLFLEIARDIOKAEPEWKI	180
QY	181	ILLRYFPVGAHESGKLGEDPKGIPNNLMPYIQQAVAGRLTELNVYGHDPYTRDGSARD	240
Db	181	VLLRYFPVGAHESGKLGEDPRGIPNNLMPYIQQAVAGRLPELVYGHDPYTRDGSARD	240
QY	241	YIHYMDADGHIARLKFTEENIGCTAYNLGTRGTSVLEVTAEKASGKKIPVKLC	300
Db	241	YIHYMDADGHIARLKFTEENIGCTAYNLGTRGSSVLEVMVAEPEKASGKKIAUKCP	300
QY	301	RRPGDATEVYASTERAKELGWKANYGVBEEMCRDQWNAKNPNPWGAGKP	350
Db	301	RRPGDATEVYASTAKAEKELGWKAKYGVBEEMCRDQWNAKNPNPWGYSGRP	350

RESULT 2

B86261
 UDPglucose 4-epimerase (EC 5.1.3.2) [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
 C:Accession: B86261
 R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B86261
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-351 <STO>
 A:Cross-references: GB:AB005172; NID:98698725; PIDN:AAF78483.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
 C:Keywords: isomerase

Query Match 85.5%; Score 1578; DB 2; Length 351;
 Best Local Similarity 85.6%; Pred. No. 8.7e-122;
 Matches 297; Conservative 16; Mismatches 34; Indels 0; Gaps 0;

```

QY 3 SSSQHILVTGGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPILLSQNLQ 62
DB 4 SVEQNILVTGGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPILLSQNLQ 63

QY 63 FTQGLNRDLEKLFSTTFDAVIHPAGLKAVAESVAKPRRYFDNVLGTINLYEFMAK 122
DB 64 FNLDLRNKGDIKLFKQRFDAVIHPAGLKAVGESVEKGRYFDNVLGTINLYEFMAK 123

QY 123 YNCKMVFSSATVYGOPEKIPCEBDFKQAMPYGRTKLFLEIARDIOKAEPEWKIL 182
DB 124 YNCKMVFSSATVYGOPEKIPCEBDFKQAMPYGRTKLFLEIARDIOKAEPEWKIL 183

QY 183 LRYFNPVGAHESGKLGPDKGIPNNLMPYIQQVAVGRLTELNVYGHDPYTRDGSARDYI 242
DB 184 LRYFNPVGAHESGKLGPDKGIPNNLMPYIQQVAVGRLTELNVYGHDPYTRDGSARDYI 243

QY 243 HVMDLADGHTAALKLFSTTNGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 302
DB 244 HVMDLADGHTAALKLFSTTNGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 303

QY 303 PGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGYAGK 349
DB 304 SGDATAVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGYAGK 350

```

RESULT 3
 S62783
 UDPglucose 4-epimerase (EC 5.1.3.2) - Arabidopsis thaliana
 N:Alternate names: uridine diphosphate glucose epimerase
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 24-Aug-1996 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
 C:Accession: S62783
 R:Doermann, P.; Benning, C.
 Arch. Biochem. Biophys. 327, 27-34, 1996
 A:Title: Functional Expression of Uridine 5'-diphospho-glucose 4-epimerase (EC 5.1.3.2)
 A:Reference number: S62783; MUID:96201343; PMID:8615692
 A:Accession: S62783
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-351 <DOE>
 A:Cross-references: EMBL:Z54214; NID:gl143391; PIDN:CAA90941.1; PID:gl143392
 C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
 C:Keywords: galactose metabolism; isomerase

F:9-347/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 85.1%; Score 1571; DB 2; Length 351;
 Best Local Similarity 85.6%; Pred. No. 3.3e-121;
 Matches 297; Conservative 16; Mismatches 34; Indels 0; Gaps 0;

```

QY 3 SSSQHILVTGGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPILLSQNLQ 62
DB 4 SVEQNILVTGGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPILLSQNLQ 63

QY 63 FTQGLNRDLEKLFSTTFDAVIHPAGLKAVAESVAKPRRYFDNVLGTINLYEFMAK 122
DB 64 FNLDLRNKGDIKLFKQRFDAVIHPAGLKAVGESVEKGRYFDNVLGTINLYEFMAK 123

QY 123 YNCKMVFSSATVYGOPEKIPCEBDFKQAMPYGRTKLFLEIARDIOKAEPEWKIL 182
DB 124 YNCKMVFSSATVYGOPEKIPCEBDFKQAMPYGRTKLFLEIARDIOKAEPEWKIL 183

QY 183 LRYFNPVGAHESGKLGPDKGIPNNLMPYIQQVAVGRLTELNVYGHDPYTRDGSARDYI 242
DB 184 LRYFNPVGAHESGKLGPDKGIPNNLMPYIQQVAVGRLTELNVYGHDPYTRDGSARDYI 243

QY 243 HVMDLADGHTAALKLFSTTNGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 302
DB 244 HVMDLADGHTAALKLFSTTNGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 303

QY 303 PGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGYAGK 349
DB 304 SGDATAVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGYAGK 350

```

RESULT 4

D96657

hypothetical protein F16M1.9 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96657

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-353 <STO>

A:Cross-references: GB:AB005173; NID:gl0092244; PIDN:AAG12659.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match 82.5%; Score 1522; DB 2; Length 353;
 Best Local Similarity 81.4%; Pred. No. 3.5e-117;
 Matches 285; Conservative 25; Mismatches 38; Indels 2; Gaps 1;

```

QY 3 SSSQHILVTGGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPILLSQNLQ 62
DB 4 SVEQNILVTGGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPILLSQNLQ 63

QY 63 FTQGLNRDLEKLFSTTFDAVIHPAGLKAVAESVAKPRRYFDNVLGTINLYEFMAK 120
DB 64 FNLDLRNKGDIKLFKQRFDAVIHPAGLKAVGESVGNPRYFDNVLGTINLYEFMAK 123

QY 121 AKYNCKMVFSSATVYGOPEKIPCEBDFKQAMPYGRTKLFLEIARDIOKAEPEWKI 180
DB 124 AKYNCKMVFSSATVYGOPEKIPCEBDFKQAMPYGRTKLFLEIARDIOKAEPEWKI 183

```

QY 181 ILLRYFNPVGAHSGKLGEDPKGIPNNLMPIYQQVAVGRLTELNVYGHDPYTRDGSALRD 240
 DB 184 ILLRYFNPVGAHSGRIGEDPKGIPNNLMPIYQQVAVGRLPELVNFGHDYPTWDGSAVRD 243
 QY 241 YHVMDLADGHIALLRKLFTEENIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 300
 DB 244 YHVMDLADGHIALLRKLFTEENIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 303
 QY 301 RRGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGAGKP 350
 DB 304 RRGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGAGKP 353

RESULT 5
 T10496
 UDPglucose 4-epimerase (EC 5.1.3.2) (clone GEPI42) - guar
 N;Alternate names: UDP-galactose 4-epimerase
 C;Species: *Cyamopsis tetragonoloba* (guar, cluster bean)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C;Accession: T10496
 R;Brunstedt, J.; Joersbo, M.; Pedersen, S.G.; Marcussen, J.
 Plant Sci. 142, 147-154, 1999
 A;Title: Isolation and expression of two cDNA clones encoding UDP-galactose epimerase ex
 A;Reference number: Z17058
 A;Accession: T10496
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-354 <BRU>
 A;Cross-references: EMBL:AJ005081
 C;Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
 C;Keywords: galactose metabolism; isomerase
 F;12-350/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 81.3%; Score 1500; DB 2; Length 354;
 Best Local Similarity 79.6%; Pred. No. 2.2e-115;
 Matches 277; Conservative 32; Mismatches 39; Indels 0; Gaps 0;

QY 2 VSSSHILVTGGAGFTGTHVTVOLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQL 61
 DB 6 MASGETILVTGGAGFTGSHVTVOLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQL 65
 QY 62 QFTQGLNRDRLLEKLFSTKTFDAVTHFAGLKAVASVAKPRYFDNLTGNTINLYEFNA 121
 DB 66 HFHGLDLNRHLDILFSKTKFDVTHFAGLKAGVSVLNPNSNYDNNLVATINLFQVMS 125
 QY 122 KYNCKKMFSSATVVGQPEKIPCEEDFKLOANPYGRKLFLEETARDIOKAEPEWKII 181
 DB 126 KFNCKKLVISSSATVVGQPDQIPCEVDSNLHAMNPYGRSKLFVEEVARDIOKAEAEWRII 185
 QY 182 LLRYFNPVGAHSGKLGEDPKGIPNNLMPIYQQVAVGRLTELNVYGHDPYTRDGSALRDY 241
 DB 186 LLRYFNPVGAHSGQIGEDPRGLPNNLMPIYQQVAVARPELVNFGHDYPTKDGTAIRDY 245
 QY 242 IHVMDLADGHIALLRKLFTEENIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 301
 DB 246 IHVMDLADGHIALLRKLFTEENIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 305
 QY 302 RRGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGAGKP 349
 DB 306 RRGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGAGKP 353

RESULT 6
 T10498
 UDPglucose 4-epimerase (EC 5.1.3.2) (clone GEPI48) - guar
 N;Alternate names: UDP-galactose 4-epimerase
 C;Species: *Cyamopsis tetragonoloba* (guar, cluster bean)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C;Accession: T10498
 R;Brunstedt, J.; Joersbo, M.; Pedersen, S.G.; Marcussen, J.
 Plant Sci. 142, 147-154, 1999
 A;Title: Isolation and expression of two cDNA clones encoding UDP-galactose epimerase ex

A;Reference number: Z17058
 A;Accession: T10498
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-350 <BRU>
 A;Cross-references: EMBL:AJ005082
 C;Function:
 A;Description: catalyzes the reversible epimerization of UDP-galactose and UDP-glucose
 A;Pathway: carbohydrate metabolism
 C;Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
 C;Keywords: galactose metabolism; isomerase
 F;6-342/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 67.5%; Score 1246; DB 2; Length 350;
 Best Local Similarity 67.5%; Pred. No. 1.6e-94;
 Matches 233; Conservative 44; Mismatches 66; Indels 2; Gaps 2;

QY 4 SSOHILVTGGAGFTGTHVTVOLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQLQF 63
 DB 2 SSOHILVTGGAGFTGSHVTVOLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQLQF 60
 QY 64 TOGDLNRDRLLEKLFSTKTFDAVTHFAGLKAVASVAKPRYFDNLTGNTINLYEFNAKY 123
 DB 61 HKLDLRDRLLEKLFSTKTFDSVTHFAGLKAVGSVQKPLLYDNNLTGTVLFEVMAAH 120
 QY 124 NCKKMFSSATVVGQPEKIPCEEDFKLOANPYGRKLFLEETARDIOKAEPEWKIILL 183
 DB 121 GCKKLVFSSSATVVGQPEKIPCEEDFKLOANPYGRKLFLEETARDIOKAEPEWKIILL 180
 QY 184 RYFNPVGAHSGKLGEDPKGIPNNLMPIYQQVAVGRLTELNVYGHDPYTRDGSALRDY 243
 DB 181 RYFNPVGAHSGQIGEDPRGLPNNLMPIYQQVAVGRLTELNVYGHDPYTRDGSALRDY 240
 QY 244 VMDLADGHIALLRKLFTEENIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 303
 DB 241 VMDLADGHIALLRKLFTEENIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 299
 QY 304 GDATVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGAG 348
 DB 300 GDATVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGAG 344

RESULT 7
 T01881
 UDPglucose 4-epimerase (EC 5.1.3.2) F8M12.10 - *Arabidopsis thaliana*
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 21-Jan-2000
 C;Accession: T01881
 R;Madsen, C.; Graves, T.; Cotton, M.; Modde, T.
 submitted to the EMBL Data Library, July 1998
 A;Description: The sequence of A. thaliana F8M12.
 A;Reference number: Z14450
 A;Accession: T01881
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-350 <MAD>
 A;Cross-references: EMBL:AF080118; NID:G3513725; PIDN:AAC33955.1; PID:G3513739
 A;Experimental source: cultivar Columbia
 C;Genetics:

A;Map position: 4
 A;Introns: 61/3; 78/2; 123/3; 157/3; 233/3; 260/1; 287/3; 318/2
 A;Note: F8M12.10
 C;Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
 C;Keywords: galactose metabolism; isomerase
 F;5-341/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 66.6%; Score 1229; DB 2; Length 350;
 Best Local Similarity 66.7%; Pred. No. 3.9e-93;
 Matches 228; Conservative 50; Mismatches 52; Indels 2; Gaps 2;

QY 5 SSOHILVTGGAGFTGTHVTVOLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQLQFT 64
 DB 2 ARNVLSGAGYIGSHVTVOLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQLQFT 60

QY 65 QGDLNRDDLEKLFSTTTFDAVIHFAGLKAVAESVAKPRRYFDENLVGTINLYEFMAKYN 124
 DB 61 QVDLRDRSALEKIFSETKFDVAIHFAGLKAVGESVEKELLYNNLVGTITLLEVAQHG 120
 QY 125 CKKQVFSSSATVYGOPEKIPCEBDFKLQAMNPGYGRTKLFLEIARDIOKABPEWKIILLR 184
 DB 121 CKNLVFGSSATVYGSPEKVPCTEETFPISALNPGYGRTKLFIEICRDVYGSPEWKIILLR 180
 QY 185 YFNPGVAGHSGKLGEDPKGIPNNLMPIYQOVAVGRLTSLNLYGHDPYPRDGSARDYIHV 244
 DB 181 YFNPGVAGHSGDGEDPGIPNNLMPIYQOVAVGRPHLTVPFGNDYNTKDTGVRDYIHV 240
 QY 245 MDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAPEKASGKKIPVKLCPRRPG 304
 DB 241 IDLADGHIAALRKLFTCK-IGCEVNLGTGNGTSVLEMTAPEKASGKKIPLVLAGRRPG 299
 QY 305 DATEVYASTERAEKELGWKANYGVEMCRDWNWAKNPNWY 346
 DB 300 DAEVYASTERAESELNWKAKYGIEMCRDLNWNASNNPYGY 341

RESULT 8

T04291

probable UDPglucose 4-epimerase (EC 5.1.3.2) - Arabidopsis thaliana

N:Alternate names: protein F25124.170

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jan-2000

C:Accession: T04291

R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15261

A:Accession: T04291

A:Molecule type: DNA

A:Residues: 1-351 <BEV>

A:Cross-references: EMBL:AL049525

A:Experimental source: cultivar Columbia; BAC clone F25124

C:Genetics:

A:Map position: 4

A:Introns: 62/3; 79/2; 124/3; 158/3; 234/3; 261/1; 288/3; 319/2

A:Note: F25124.170

C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

C:Keywords: isomerase

F:6-342/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 66.6%; Score 1229; DB 2; Length 351;
 Best Local Similarity 66.7%; Pred. No. 3.9e-93; Indels 2; Gaps 2;
 Matches 228; Conservative 50; Mismatches 62;

QY 5 SQHILVTGGAGFTGTHVTVQLLKAGFSVSIIDFNDSVMEAVDRVQVVGPLLSONLOFT 64
 DB 3 ARNLVSGGAGYIGSTHTVLQLLGGYVVVDNLNDSAVSLQKVKKLAAE-HGERLSGH 61
 QY 65 QGDLNRDDLEKLFSTTTFDAVIHFAGLKAVAESVAKPRRYFDENLVGTINLYEFMAKYN 124
 DB 62 QVDLRDRSALEKIFSETKFDVAIHFAGLKAVGESVEKELLYNNLVGTITLLEVAQHG 121
 QY 125 CKKQVFSSSATVYGOPEKIPCEBDFKLQAMNPGYGRTKLFLEIARDIOKABPEWKIILLR 184
 DB 122 CKNLVFGSSATVYGSPEKVPCTEETFPISALNPGYGRTKLFIEICRDVYGSPEWKIILLR 181
 QY 185 YFNPGVAGHSGKLGEDPKGIPNNLMPIYQOVAVGRLTSLNLYGHDPYPRDGSARDYIHV 244
 DB 182 YFNPGVAGHSGDGEDPGIPNNLMPIYQOVAVGRPHLTVPFGNDYNTKDTGVRDYIHV 241
 QY 245 MDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAPEKASGKKIPVKLCPRRPG 304
 DB 242 IDLADGHIAALRKLFTCK-IGCEVNLGTGNGTSVLEMTAPEKASGKKIPLVLAGRRPG 300

QY 305 DATEVYASTERAEKELGWKANYGVEMCRDWNWAKNPNWY 346

DB 301 DAEVYASTERAESELNWKAKYGIEMCRDLNWNASNNPYGY 342

RESULT 9

T08911

UDPglucose 4-epimerase (EC 5.1.3.2) T32A16.90 - Arabidopsis thaliana

N:Alternate names: protein T32A16.90

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jan-2000

C:Accession: T08911

R:Bevan, M.; Zimmermann, W.; Gruenisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Ma

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16518

A:Accession: T08911

A:Molecule type: DNA

A:Residues: 1-350 <BEV>

A:Cross-references: EMBL:AL078468; GSPDB:GN00062; ATSP:T32A16.90

A:Experimental source: cultivar Columbia; BAC clone T32A16

C:Genetics:

A:Gene: ATSP:T32A16.90

A:Map position: 4

A:Introns: 61/3; 78/2; 123/3; 157/3; 233/3; 260/1; 287/3; 318/2

C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

C:Keywords: galactose metabolism; isomerase

F:5-341/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 65.5%; Score 1208; DB 2; Length 350;

Best Local Similarity 65.5%; Pred. No. 2.1e-91;

Matches 224; Conservative 49; Mismatches 67; Indels 2; Gaps 2;

QY 5 SQHILVTGGAGFTGTHVTVQLLKAGFSVSIIDFNDSVMEAVDRVQVVGPLLSONLOFT 64

DB 2 AKSLVTGGAGYIGSTHTVLQLLGGYSAVVVDNDSAAQLQKVKKLAGE-NGNLSLPH 60

QY 65 QGDLNRDDLEKLFSTTTFDAVIHFAGLKAVAESVAKPRRYFDENLVGTINLYEFMAKYN 124

DB 61 QVDLRDRSALEKIFSETKFDVAIHFAGLKAVGESVEKELLYNNLVGTITLLEVAQHG 120

QY 125 CKKQVFSSSATVYGOPEKIPCEBDFKLQAMNPGYGRTKLFLEIARDIOKABPEWKIILLR 184

DB 121 CKNLVFGSSATVYGSPEKVPCTEETFPISALNPGYGRTKLFIEICRDVYGSPEWKIILLR 180

QY 185 YFNPGVAGHSGKLGEDPKGIPNNLMPIYQOVAVGRLTSLNLYGHDPYPRDGSARDYIHV 244

DB 181 YFNPGVAGHSGYIGEDPLGVFNLMPIYQOVAVGRPHLTVPFGDTYKDKGTGVRDYIHV 240

QY 245 MDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAPEKASGKKIPVKLCPRRPG 304

DB 241 MDLADGHIAALRKLDDLK-ISCEVNLGTGNGTSVLEMTAPEKASGKKIPLVLAGRRPG 299

QY 305 DATEVYASTERAEKELGWKANYGVEMCRDWNWAKNPNWY 346

DB 300 DAEVYASTERAEKELNWKAKYGIEMCRDLNWNASNNPYGY 341

RESULT 10

D69628

UDPglucose 4-epimerase (EC 5.1.3.2) - Bacillus subtilis

N:Alternate names: UDPgalactose 4-epimerase

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: D69628

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Ertting, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

Tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningssein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon

A:Authors: Schleich, S.; Schroeter, R.; Scofield, P.; Sekiguchi, J.; Sekowska, A.; Sero

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Torato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: D69628
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

A;Residues: 1-339 <KUN>
A;Cross-references: GB:299123; GB:AL009126; NID:G2636240; PIDN:CA15912.1; PID:G2636421
A;Experimental source: strain 168
C;Genetics:

A;Gene: gale
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C;Keywords: galactose metabolism; isomerase
F;3-335/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 56.3%; Score 1038; DB 2; Length 339;
Best Local Similarity 56.6%; Pred. No. 1.8e-77;
Matches 192; Conservative 58; Mismatches 83; Indels 6; Gaps 2;

QY 8 ILVTGGAGTGTHTVQVLLKAGSVSIIDNFNSVMEAVDRVQVGPILLSQNLQFTQGD 67
DB 3 ILVTGGAGTGTHTVQVLLKAGSVSIIDNFNSVMEAVDRVQVGPILLSQNLQFTQGD 58

QY 68 LRNRDLEKLFSTTTDAVTHFAGLKAVASVAKPRRYFDNLTGVTINLYEFMAKYNCKK 127
DB 59 LLDRBAVDSVFAENEIEAVTHFAGLKAVGESVAIPLYKYYHNNLTGTFILCEAMEKYGK 118

QY 128 MVFSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQABPEWKIILLRYFN 187
DB 119 IVFSSATVYGPETSPITEDFPGLGATNPYQGTQKLEQLRLDLHTADNENSVALLRYFN 178

QY 188 PVGAHESGKLGEDPKGIPNNLMPIYQVAVGRUETELNVYGHDPYTRDGSAIRDIHVMDL 247
DB 179 PFGAHPGSRIGEDPNIIPNNLMPIYQVAVGRUETELNVYGHDPYTRDGSAIRDIHVMDL 238

QY 248 ADGHIAALRKLFTEENIGCTAYNLGTGRGTSVLEMTAFKASGKIPVKLCPRRPGDAT 307
DB 239 AEGHVKALEKVL--NSTGADAVNLGTGTSVLEMTAFKASGKIPVKLCPRRPGDAT 296

QY 308 EVYASTERAEKELGWKANYGVEMCRDQWNAKNPMGY 346
DB 297 TCPADPAKAKRELGWKANYGVEMCRDQWNAKNPMGY 335

RESULT 11
F82419
UDP-glucose 4-epimerase VCA0774 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: F82419
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: F82419

A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-338 <HEI>
A;Cross-references: GB:AE004405; GB:AE003853; NID:G9658186; PIDN:AAF96672.1; GSPDB:GN001

A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:

A;Map position: 2
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match 53.3%; Score 983; DB 2; Length 338;
Best Local Similarity 54.9%; Pred. No. 5.8e-73;

Matches 186; Conservative 54; Mismatches 93; Indels 6; Gaps 2;
QY 8 ILVTGGAGTGTHTVQVLLKAGSVSIIDNFNSVMEAVDRVQVGPILLSQNLQFTQGD 67

DB 3 VLVTGGMGYIGTSHTCIQMIQAGMTPVILNDLNSKVTVLDRIEKVIG----VRPQFVQGD 58
QY 68 LRNRDLEKLFSTTTDAVTHFAGLKAVASVAKPRRYFDNLTGVTINLYEFMAKYNCKK 127

DB 59 IRDKALLVDLMQOHNEIAVTHFAGLKAVGESVQKPLEYDNNVNGTLVLAAMEAGVKS 118
QY 128 MVFSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQABPEWKIILLRYFN 187

DB 119 LVFSSATVYGPETSPITEDFPGLGATNPYQGTQKLEQLRLDLHTADNENSVALLRYFN 178
QY 188 PVGAHESGKLGEDPKGIPNNLMPIYQVAVGRUETELNVYGHDPYTRDGSAIRDIHVMDL 247

DB 179 PFGAHPGSRIGEDPNIIPNNLMPIYQVAVGRUETELNVYGHDPYTRDGSAIRDIHVMDL 238
QY 248 ADGHIAALRKLFTEENIGCTAYNLGTGRGTSVLEMTAFKASGKIPVKLCPRRPGDAT 307

DB 239 ADGHIAALKKVGTC--AGLHIYNLTGTSVLEMTAFKASGKIPVKLCPRRPGDAT 296
QY 308 EVYASTERAEKELGWKANYGVEMCRDQWNAKNPMGY 346

DB 297 EYADPTKAAQDLGWKATRNLTHTMAQAWCQSNPNPGY 335

RESULT 12
AC3367

UDPglucose 4-epimerase (EC 5.1.3.2) [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002

C;Accession: AC3367
R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess,

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*

A;Reference number: AD3252; PMID:11756688
A;Accession: AC3367

A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-335 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL52102.1; PID:g17982874; GSPDB:GN00190

A;Experimental source: strain 16M
C;Genetics:

A;Gene: BMEI0921
A;Map position: 1

C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C;Keywords: isomerase

Query Match 52.7%; Score 971.5; DB 2; Length 335;
Best Local Similarity 54.6%; Pred. No. 5.1e-72;

Matches 185; Conservative 53; Mismatches 94; Indels 7; Gaps 2;
QY 8 ILVTGGAGTGTHTVQVLLKAGSVSIIDNFNSVMEAVDRVQVGPILLSQNLQFTQGD 67

DB 3 ILVTGGAGTGTHTVQVLLKAGSVSIIDNFNSVMEAVDRVQVGPILLSQNLQFTQGD 58

QY 68 LRNRDLEKLFSTTTDAVTHFAGLKAVASVAKPRRYFDNLTGVTINLYEFMAKYNCKK 127
DB 59 IRDRALMEQIMIKHCKTAVTHFAGLKAVGESKPLLYDCNLTGLRLQLQAMEATGVKK 118

QY 128 MVFSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQABPEWKIILLRYFN 187
DB 119 LVFSSATVYGPDPKLPITDQPLSATNPYGRTKLVIEDMLRLDLYNSDMSWAIAILRYFN 178

QY 188 PVGAHESGKLGEDPKGIPNNLMPIYQVAVGRUETELNVYGHDPYTRDGSAIRDIHVMDL 247
DB 179 PVGAHESGLTGEDPKGIPNNLMPIYQVAVGRUETELNVYGHDPYTRDGSAIRDIHVMDL 238

QY 248 ADGHIAALRKLFTEENIGCTAYNLGTGRGTSVLEMTAFKASGKIPVKLCPRRPGDAT 307
DB 239 AAGHLKALKKL---DKPKCFANVLTGGYSLVDVIKAFEHVSNREIKYEIAPERPDVA 295

QY 308 EVYASTERAEKELGWKANYGVEMCRDQWNAKNPMGY 346
DB 296 EYADPTGFAKFLGWSAEKLNREMCODMWNQSNPNGY 334


```

A;Gene: galE
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C;Keywords: isomerase

Query Match      52.2%; Score 964; DB 2; Length 339;
Best Local Similarity 51.9%; Pred. No. 2.1e-71;
Matches 177; Conservative 59; Mismatches 99; Indels 6; Gaps 2;

QY 6 QHILVTGGAGFICTHTVWQLLKAGFSVSIIDNFDSVMEAVDRVQVGVGPLLSQLQFTQ 65
DB 3 EKILVTGGAGFICTHTVWQLLKAGFSVSIIDNFDSVMEAVDRVQVGVGPLLSQLQFTQ 65
QY 66 GDLNRDDEKLFPSKTTFDVAIHFAGLKAVAESVAKPRYFDNVLVGTINLYEFMAKYN 125
DB 59 ADIRDYDTRDIFKQEEPTGVHFAGLKAVGESTRIPLAYVDNNIAGTVSLLKAMEENN 118
QY 126 KMWFFSSATVYQPEKIPCEEDFKLOANVPYGRTKLFLEIARDIQKAPPEWKIILRY 185
DB 119 KNIFSSATVYGDPTVPIEDPPLSVNTNPGYGRTKLMLEELTDIYKADSEWNVLLRY 178
QY 186 FNPVGAHESKLGEDPKGIPNNLMPYIQVAVGRLTELNVYGHDPYTRDGSALRDIHVM 245
DB 179 FNPVGAHESKLGEDPKGIPNNLMPYIQVAVGRLTELNVYGHDPYTRDGSALRDIHVM 238
QY 246 DLADGHIAALRKLTFTENICTAYNLGTRGTSVLEMTAFKASGKKIPVKLCPRRPGD 305
DB 239 DLAKGHVAALKKI--QKSGNLVNYLGTGKYSVLEIIQNMKAVGCRPIRYIVERRPGD 296
QY 306 ATEVYASTERAELGKWKANYGVEEMCRDOWNAKNNPMGY 346
DB 297 IAACYSDDPAKAKAELGWEAELDTQMCDAWRQSKHPNGF 337

RESULT 15
T19989
hypothetical protein C47B2.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
R;Kershaw, J.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z19208
A;Accession: T19989
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-347 <WLL>
A;Cross-references: EMBL:Z99709; PIDN: CAB16861.1; GSPDB: GN00019; CESP: C47B2.6
A;Experimental source: clone C47B2
C;Genetics:
A;Gene: CESP:C47B2.6
A;Map position: 1
A;Introns: 39/1; 115/3; 258/3; 296/3
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match      52.1%; Score 962; DB 2; Length 347;
Best Local Similarity 52.1%; Pred. No. 3.2e-71;
Matches 183; Conservative 59; Mismatches 94; Indels 16; Gaps 4;

QY 7 HILVTGGAGFICTHTVWQLLKAGFSVSIIDNFDSVMEAVDRVQVGVGPLLSQLQ 60
DB 2 HILVTGAAGFISGHTVLELNSGYTVLCIDNFANISDEHGNALSKRVAQLTG----KD 57
QY 61 LQFTQGLDRNRDDEKLFPSKTTFDVAIHFAGLKAVAESVAKPRYFDNVLVGTINLYEFM 120
DB 58 VPFQNVDCDEAALEKVFSENKFDGIHLLAALKAVGESVAKPQYYSNNLVASLNTQMC 117
QY 121 AKYCKKQWVSSATVYQPEKIPCEEDFKLOA----MNPYGRTKLFLEIARDIQKAP 176
DB 118 LKYNKQWVSSATVYQPEKIPCEEDFKLOA----MNPYGRTKLFLEIARDIQKAP 175
QY 177 EWKITLLRYNPNVGAHESKLGEDPKGIPNNLMPYIQVAVGRLTELNVYGHDPYTRDGS 236
DB 176 EWNVLLRYNPNVGAHESKLGEDPKGIPNNLMPYIQVAVGRLTELNVYGHDPYTRDGS 235

A;Gene: galE
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C;Keywords: isomerase

Query Match      52.4%; Score 966; DB 2; Length 339;
Best Local Similarity 51.9%; Pred. No. 1.5e-71;
Matches 177; Conservative 59; Mismatches 99; Indels 6; Gaps 2;

QY 6 QHILVTGGAGFICTHTVWQLLKAGFSVSIIDNFDSVMEAVDRVQVGVGPLLSQLQFTQ 65
DB 3 EKILVTGGAGFICTHTVWQLLKAGFSVSIIDNFDSVMEAVDRVQVGVGPLLSQLQFTQ 65
QY 66 GDLNRDDEKLFPSKTTFDVAIHFAGLKAVAESVAKPRYFDNVLVGTINLYEFMAKYN 125
DB 59 ADIRDYDTRDIFKQEEPTGVHFAGLKAVGESTRIPLAYVDNNIAGTVSLLKAMEENN 118
QY 126 KMWFFSSATVYQPEKIPCEEDFKLOANVPYGRTKLFLEIARDIQKAPPEWKIILRY 185
DB 119 KNIFSSATVYGDPTVPIEDPPLSVNTNPGYGRTKLMLEELTDIYKADSEWNVLLRY 178
QY 186 FNPVGAHESKLGEDPKGIPNNLMPYIQVAVGRLTELNVYGHDPYTRDGSALRDIHVM 245
DB 179 FNPVGAHESKLGEDPKGIPNNLMPYIQVAVGRLTELNVYGHDPYTRDGSALRDIHVM 238
QY 246 DLADGHIAALRKLTFTENICTAYNLGTRGTSVLEMTAFKASGKKIPVKLCPRRPGD 305
DB 239 DLAKGHVAALKKI--QKSGNLVNYLGTGKYSVLEIIQNMKAVGCRPIRYIVERRPGD 296
QY 306 ATEVYASTERAELGKWKANYGVEEMCRDOWNAKNNPMGY 346
DB 297 IAACYSDDPAKAKAELGWEAELDTQMCDAWRQSKHPNGF 337

RESULT 14
C98054
UDPglucose 4-epimerase (EC 5.1.3.2) [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: C98054
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID: 21429245; PMID: 11544234
A;Accession: C98054
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-339 <KUR>
A;Cross-references: GB:AE007317; PIDN: AAL00264.1; PID: g15459117; GSPDB: GN00174
C;Genetics:
```

Qy	237	AIRDYIHVMDLADCHIAALRKLFPTTENIGCTAINLGTGRGTSVLEWVAFEXASGKKIPV	296
Db	236	GVEDYIHVVDLAKGHVKAFDRIKTVGNIGTEIYNLGTGVGYSVRQWVDALKKVSGRDIPV	295
Qy	297	KLCPRRPGDATEVYASTERAELGHWKANYGVEMCRDOWNWAKNPNWGYA	347
Db	296	KIGVPRPGDVASYCDPSLAQEXLGWRAETGLEEMCADLWNQTKNPOGFS	346

Search completed: March 22, 2004, 21:56:00
Job time : 28 secs

Db 1 MVASSQKILVTGAGFIGHTVTVQLLNGFNVSIIIDNFDSNMEAVREVVWGSNLSON 60
 QY 61 LQFTQGLNRDDEKLEKSTTFDAVIFHAGLKAVASVAKPRYFDNLTGTLNLYEFM 120
 Db 61 LEFTGLDRLKDDLEKLEKSKFDVIFHAGLKAVGESVENPRYFDNLTGTLNLYEFM 120
 QY 121 AKYNCKMVFSSSATVYQGPKEIPCEBDFKLQANMPYGRTKLFLEETARDIOKAEPEWKI 180
 Db 121 AKHNCKMVFSSSATVYQGPKEIPCEBDFKLQANMPYGRTKLFLEETARDIOKAEPEWKI 180
 QY 181 ILLRYFNPVGAHESKLEKDEPKGIENMLPVIQQVAVGRLTAVNLYGHDPYTRDGSARD 240
 Db 181 VLLRYFNPVGAHESKLEKDEPKGIENMLPVIQQVAVGRLTAVNLYGHDPYTRDGSARD 240
 QY 241 YHVNMDLADGHIAALRLFTTENICTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 300
 Db 241 YHVNMDLADGHIAALRLFTTENICTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 300
 QY 301 RPPGDATEVYASTAEKELGWKANYGVEMCRDQNNWAKNPNWGYAGKP 350
 Db 301 RPPGDATEVYASTAEKELGWKANYGVEMCRDQNNWAKNPNWGYAGKP 350

RESULT 2

GAEL ARATH STANDARD; PRT; 351 AA.
 AC Q42605, Q9LPX1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 DE galactose 4-epimerase).
 GN AT1G12780 OR F13K23.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=cv Columbia;
 RX MEDLINE=96201343; PubMed=8615692;
 RA Doermann P., Benning C.;
 RT "Functional expression of uridine 5'-diphospho-glucose 4-epimerase
 RT (EC 5.1.3.2) from Arabidopsis thaliana in Saccharomyces cerevisiae and
 RT Escherichia coli.";
 RL Arch. Biochem. Biophys. 327:27-34 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Ataydo R., Bowman C.L., Brooks S.V.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Millschler J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RT Nature 408:816-820 (2000).
 CC -|- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -|- COFACTOR: Does not seem to necessitate NAD for activity.
 CC -|- PATHWAY: Galactose metabolism; third step.

CC -|- PATHWAY: Biosynthesis of UDP-galactose as precursor of
 CC galactolipids and cell wall polysaccharides.
 CC -|- TISSUE SPECIFICITY: Ubiquitously expressed. Highest expression
 CC in stems and roots.
 CC -|- MISCELLANEOUS: Shows a broad pH optimum of 7.0 to 9.5 and a Km of
 CC 0.11 mM.
 CC -|- SIMILARITY: Belongs to the sugar epimerase family.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z54214; CAA90941.1; -
 CC EMBL; AC012187; AAF78483.1; -
 CC PIR; B62621; B62621.
 CC PIR; S62783; S62783.
 CC HSP; P09147; IXEL.
 CC InterPro; IPR001509; Epimerase_Dh.
 CC InterPro; IPR005886; GalE.
 CC InterPro; IPR000205; NAD BS.
 CC InterPro; IPR008089; Nuc_sugar_epim.
 CC Pfam; PF01370; Epimerase_1.
 CC PRINTS; PR01713; NUCBPIMERASE.
 CC TIGRfams; TIGR01179; GalE; 1.
 KW isomerase; NAD; Galactose metabolism; Multigene family.
 FT NP BIND 8 39 NAD (POTENTIAL).
 FT CONFLICT 102 103 NP -> KG (IN REF. 1).
 FT CONFLICT 341 341 N -> F (IN REF. 1).
 SQ SEQUENCE 351 AA; 39157 MW; FDB640FB26A1DD1D CRC64;
 Query Match 85.5%; Score 1578; DB 1; Length 351;
 Best Local Similarity 85.6%; Pred. No. 3e-118;
 Matches 297; Conservative 16; Mismatches 34; Indels 0; Gaps 0;
 QY 3 SSSQHLVTGGAGFIGHTVTVQLLNGFVSIIIDNFDSNMEAVREVVWGSNLSONIQ 62
 Db 4 SVEQNLTVTGGAGFIGHTVTVQLLNGFVKVSIIDNFDSNMEAVREVVWGSNLSONIQ 63
 QY 63 FTQGLNRDDEKLEKSTTFDAVIFHAGLKAVASVAKPRYFDNLTGTLNLYEFMAK 122
 Db 64 FNLGDLNRKGDIEKLFKQRFDAVIFHAGLKAVGESVENPRYFDNLTGTLNLYETMAK 123
 QY 123 YNCKMVFSSSATVYQGPKEIPCEBDFKLQANMPYGRTKLFLEETARDIOKAEPEWKIL 182
 Db 124 YNCKMVFSSSATVYQGPKEIPCEBDFKLQANMPYGRTKLFLEETARDIOKAEPEWKIL 183
 QY 183 LRYFNPVGAHESKLEKDEPKGIENMLPVIQQVAVGRLTAVNLYGHDPYTRDGSARDYI 242
 Db 184 LRYFNPVGAHESKLEKDEPKGIENMLPVIQQVAVGRLTAVNLYGHDPYTRDGSARDYI 243
 QY 243 HVMDLADGHIAALRLFTTENICTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 302
 Db 244 HVMDLADGHIAALRLFTTENICTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 303
 QY 303 PGDATEVYASTAEKELGWKANYGVEMCRDQNNWAKNPNWGYAGK 349
 Db 304 SGDATEVYASTAEKELGWKANYGVEMCRDQNNWAKNPNWGYQNK 350
 RESULT 3
 ID GAEL CYATE STANDARD; PRT; 354 AA.
 AC O65780;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase GEPI42 (EC 5.1.3.2) (Galactowaldenase) (UDP-
 DE galactose 4-epimerase).
 OS Cyamopsis tetragonoloba (Guar) (Cluster bean).


```
-----
CC  EMBL; L41668; AAC86498.1; -.
DR  EMBL; AF022382; AAC39645.1; -.
DR  EMBL; AL031295; CAB40159.1; -.
DR  EMBL; BC001273; AAH01273.1; -.
DR  EMBL; BC050685; AAH050685.2; -.
DR  PDB; 1EK5; 17-MAY-00.
DR  PDB; 1EK6; 17-MAY-00.
DR  PDB; 1H2J; 09-MAY-01.
DR  PDB; 1I3K; 20-JUN-01.
DR  PDB; 1I3L; 20-JUN-01.
DR  PDB; 1I3M; 20-JUN-01.
DR  PDB; 1I3N; 20-JUN-01.
DR  Genew; HGNC:4116; GALE.
DR  GK; Q14376; -.
DR  MIM; 606953; -.
DR  MIM; 230350; -.
DR  InterPro; IPR001509; Epimerase_Dh.
DR  InterPro; IPR005886; GALE.
DR  InterPro; IPR008099; Nuc_sugar_epim.
DR  Pfam; PF01370; Epimerase; 1.
DR  PRINTS; PR01713; NUCEPIMERASE.
DR  TIGRFAMs; TIGR01179; gale; 1.
DR  Isomerase; NAD; Galactose metabolism; Disease mutation; Polymorphism;
KW  3D-structure.
FT  NP_BIND. 4 35 NAD (POTENTIAL).
FT  ACT_SITE 157 157 BASE.
FT  VARIANT 34 34 N -> S (in EDG; peripheral; 70% of wild-
type activity).
FT  VARIANT 90 90 /FTid=VAR_002539.
FT  VARIANT 94 94 G -> E (in EDG).
FT  VARIANT 94 94 /FTid=VAR_002540.
FT  VARIANT 94 94 V -> M (in EDG; 5% of normal activity
with respect to UDP-galactose; 24% of
normal activity with respect to UDP-N-
acetyl-galactosamine).
FT  VARIANT 103 103 /FTid=VAR_010058.
FT  VARIANT 103 103 D -> G (in EDG).
FT  VARIANT 180 180 /FTid=VAR_002541.
FT  VARIANT 180 180 V -> A.
FT  VARIANT 183 183 /FTid=VAR_002542.
FT  VARIANT 183 183 L -> P (in EDG; peripheral; 4% of wild-
type activity).
FT  VARIANT 257 257 /FTid=VAR_002543.
FT  VARIANT 313 313 K -> R (in EDG).
FT  VARIANT 313 313 L -> M (in EDG).
FT  VARIANT 319 319 /FTid=VAR_002545.
FT  VARIANT 319 319 G -> E (in EDG).
FT  SEQUENCE 348 AA; 38310 MW; DAA0C0C0C378CCEB CRC64;
Query Match 53.5%; Score 986.5; DB 1; Length 348;
Best Local Similarity 54.2%; Pred. No. 3.6e-71;
Matches 189; Conservative 59; Mismatches 88; Indels 13; Gaps 4;
QY 5 SQHILVTGGAGFGTGTHTVQLLKAGFSVSIIDNFN-----SVMEAVDRVQVVGPLLS 58
DB 2 AEKVLVTGGAGYIGSHTVLELEAGYLPVVIDNFHNAFRGGSLPESLARVQELTG---- 57
QY 59 QNLQFTQGLNRDDLEKLFSTKTFDAVHFAGLKAVASVAKPRYFDFNLVGTINLYE 118
DB 58 RSVEFEEMDILDQALQRLFKKYSFMAVHFAGLKAVGSVQKPLDYRNVLTGTIQLLE 117
QY 119 FMAKYNCKKGVFSSSATVYQPEKIPCEEDFKL-QAMNPGYGRTKLFLERDIARDIOKAEPE 177
DB 118 IMKAHGVNVLVFSSATVYGNPQYLPDEAHPGTGCTNPGYKSKFFIEMIRDLQADKT 177
QY 178 WKIILLRYNPVGAHSGKLGEDPKGIPNNILPYIOOVAVGRLTETLVYCHYFTRDGA 237
DB 178 MNVLLRYNFTGAHSGCIGEDPQGI PNNLAPYVSQVAIGREALNVFGNDYDTEGTG 237
QY 238 IRDYIHVMDLADGHIAALRKLFTTENIGTAVNLGTGRGTSVLEMTAFKASGKIPVK 297
-----
```

```
-----
DB 238 VRDYIHVVDLAKGHIAALRKL--KEQCGCRITYNLGTGTGYSLQWQVAMERKASGKIPYK 295
QY 298 LCPRRPGDATEVYASTERAELGKWKANYGVVEENCRDQWNAKNPNWGY 346
DB 296 VVARREGDVAACYANPFLAQLAEELGWTAAALGLDRMCEDLWRMOKNPSPGF 344
RESULT 10
GALE_RAT STANDARD; PRT; 347 AA.
AC P18645;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
galactose 4-epimerase).
GN GALE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90384840; PubMed=2205840;
RA Zeschmigg M., Wilcken-Bergmann B., Starzinski-Powitz A.;
RT "cDNA from rat cells with reconstitutive galactose-epimerase activity
in B. coli.";
RL Nucleic Acids Res. 18:5289-5289(1990).
CC -!- FUNCTION: Catalyzes two distinct but analogous reactions: the
epimerization of UDP-glucose to UDP-galactose and the
epimerization of UDP-N-acetylglucosamine to UDP-N-
acetyl-galactosamine.
CC -!- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: Galactose metabolism; third step.
CC -!- SIMILARITY: Belongs to the sugar epimerase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; X53949; CAA37897.1; -.
DR PIR; S11223; S11223.
DR HSSP; P09147; 1XEL.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; GALE.
DR InterPro; IPR008099; Nuc_sugar_epim.
DR Pfam; PF01370; Epimerase_1.
DR PRINTS; PR01713; NUCEPIMERASE.
DR TIGRFAMs; TIGR01179; gale; 1.
KW Isomerase; NAD; Galactose metabolism.
FT NP_BIND 4 35 NAD (POTENTIAL).
FT ACT_SITE 156 156 BASE (BY SIMILARITY).
FT SEQUENCE 347 AA; 38225 MW; 625B81546E699143 CRC64;
Query Match 51.1%; Score 943; DB 1; Length 347;
Best Local Similarity 52.9%; Pred. No. 1e-67;
Matches 183; Conservative 53; Mismatches 98; Indels 12; Gaps 3;
QY 6 QHILVTGGAGFGTGTHTVQLLKAGFSVSIIDNFNSVM-----EAVDRVQVVGPLLSQ 59
DB 3 EKVLVTGGAGYIGSHTVLELEAGYSPVIDNFHNSIRGSDSNPESLRVQELTG----R 58
QY 60 NLQFTQGLNRDDLEKLFSTKTFDAVHFAGLKAVASVAKPRYFDFNLVGTINLYE 119
DB 59 SVFEFEEMDILDQALQRLFKKYNFKAIVHFAGLKAVGSVQKPLDYRNVLTGTIQLLEI 118
-----
```

QY 120 MAKYNCKMVFSSSATYVGOPEKIPCEEDFKLQAMNYPYGTKLFLBEIARDIOKASPEWK 179
 Db 119 MRAMGVSLVFSSSATYVGPVPSASGRPPHRCGTPYKSKFFIEMIQDLCLADTAWN 178
 QY 180 IILLRYENPNVGAHESKGLGDPKGI PNNLMPIYQOAVAGRLTELNVYGHDPYPRDGSAIR 239
 Db 179 AVLLRYPIPIGAHRSARIGEDPGQIPNNLMPIYVQVAGREALNVFGDDYATEDGTGYR 238
 QY 240 DYIHMVLAAGHIAALRKLFTTNNIGCTANLGTGTGTSVLEMTAFKASGKKIPVKLC 299
 Db 239 DYIHMVLAAGHIAALRKL--KEQCGCRIYNLGTGTGTSVLEMTAFKASGKKIPYKVV 296
 QY 300 PRPGDATEVYASTERAELKGMKANYGVEMCRDQNNWAKNPNWG 345
 Db 297 ARREGDVAAACYPNSLAHEELGWTALGLDRNCEDLWRWQKQNPWG 342

RESULT 11
 GALE NEIMA
 ID _GALE NEIMA STANDARD; PRT; 339 AA.
 AC P5697;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 18-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 galactose 4-epimerase).
 GN GALE OR NMA0203.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
 RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis Z2491.";
 RL Nature 404:502-506(2000).
 CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE IN THE INCORPORATION OF
 GALACTOSE INTO MENINGOCOCCAL LIPOPOLYSACCHARIDE SURFACE MOLECULES,
 WHICH ARE IMPORTANT FOR PATHOGENESIS.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -!- COFACTOR: NAD.
 CC -!- PATHWAY: Galactose metabolism; third step.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SIMILARITY: Belongs to the sugar epimerase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL162752; CAB83517.1; -
 DR EMBL; F82014; F82014.
 DR HSPF; P09147; IXL.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GALE.
 DR Pfam; PF01370; Epimerase; 1.
 DR TIGRFAMs; TIGR01179; gale; 1.
 KW Isomerase; NAD; Galactose metabolism; Complete proteome.
 FT NP_BIND 339 34
 SQ SEQUENCE 339 AA; 37066 MW; A68DB884417B0C1C CRC64;

Query Match 50.3%; Score 928.5; DB 1; Length 339;

Best Local Similarity 52.5%; Pred. No. 1.4e-66;
 Matches 180; Conservative 58; Mismatches 96; Indels 9; Gaps 4;

QY 6 OHILVTGAGTGTHTVYVOLLKAGFSVSIIDFNDSVMEAVDRVQVVGPLLSQNLQFTQ 65
 Db 2 KNILVTGAGTGTHTVYVOLLKAGFSVSIIDFNDSVMEAVDRVQVVGPLLSQNLQFTQ 57
 QY 66 GDLNRDDLEKFLSKTTFDAVHFAGLKAFAESVAKPRYDFNLTGVTINLYEFNAKNC 125
 Db 58 GDIRELRLRIFAENRIDSVIHFAGLKAFAESVAKPRYDFNLTGVTINLYEFNAKNC 117
 QY 126 KMVPSSSATYVGOPEKIPCEEDFKL-QAMNYPYGTKLFLBEIARDIOKASPEWKIILLR 184
 Db 118 PSIVPSSSATYVGOPEKIPCEEDFKL-QAMNYPYGTKLFLBEIARDIOKASPEWKIILLR 177
 QY 185 YFNPVGAHESKGLGDPKGI PNNLMPIYQOAVAGRLTELNVYGHDPYPRDGSAIRDYIHV 244
 Db 178 YFNPVGAHESKGLGDPKGI PNNLMPIYQOAVAGRLTELNVYGHDPYPRDGSAIRDYIHV 237
 QY 245 MDLADGHIAALRKLFTTNNI-GCTAYNLGTGRTSVLEMTAFKASGKKIPVKLCPRRP 303
 Db 238 MDLADGHIAALRKLFTTNNI-GCTAYNLGTGRTSVLEMTAFKASGKKIPVKLCPRRP 294
 QY 304 GDATEVYASTERAELKGMKANYGVEMCRDQNNWAKNPNWG 346
 Db 295 GDLACFYADPSYTKAQIGWQTDLAQWEDSWRWSNPNNGY 337

RESULT 12
 GALE NEIMC
 ID _GALE NEIMC STANDARD; PRT; 339 AA.
 AC P56986; Q59617; Q59624;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 galactose 4-epimerase).
 GN GALE.
 OS Neisseria meningitidis (serogroup C).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=135720;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FAM20 / Serogroup C;
 RX MEDLINE=95310004; PubMed=7790063;
 RA Lee F.K., Stephens D.S., Gibson B.W., Engstrom J.J., Zhou D.,
 RA Apicella M.A.;
 RT "Microheterogeneity of Neisseria lipooligosaccharide: analysis of a
 UDP-glucose 4-epimerase mutant of Neisseria meningitidis NMB.";
 RL Infect. Immun. 63:2508-2515(1995).
 CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE IN THE INCORPORATION OF
 GALACTOSE INTO MENINGOCOCCAL LIPOPOLYSACCHARIDE SURFACE MOLECULES,
 WHICH ARE IMPORTANT FOR PATHOGENESIS.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -!- COFACTOR: NAD.
 CC -!- PATHWAY: Galactose metabolism; third step.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SIMILARITY: Belongs to the sugar epimerase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U19595; AAA86716.1; -
 DR HSPF; P09147; IXL.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GALE.
 DR Pfam; PF01370; Epimerase; 1.

```

DR  TIGRFAMS; TIGR01179; gale; 1.
KW  isomerase; NAD; Galactose metabolism.
FT  NP BIND 3 34 NAD (POTENTIAL).
SQ  SEQUENCE 339 AA; 37030 MW; 7DC4E4ACBE046397 CRC64;

Query Match
Best Local Similarity 50.3%; Score 928.5; DB 1; Length 339;
Matches 180; Conservative 57; Mismatches 95; Indels 9; Gaps 4;

QY  8 ILVTGGAGFTGTHVQLLKAGFSVSIIDNFNSVMEAVDRVQVQVGLLSQNLQFTQGD 67
DB  4 ILVTGGFTGTHVQLLKAGFSVSIIDNFNSVMEAVDRVQVQVGLLSQNLQFTQGD 59
QY  68 LNRDDLEKLFSTTDFAVIHFAHAGLKAFAESVAKPRYDFNLTGNTINLYEFMAKYNCKK 127
DB  60 IRDREILRRIFAENRDSVHFAGLKAFAESVAKPRYDFNLTGNTINLYEFMAKYNCKK 119
QY  128 MVFSSSATVYQPEKIPCEEDFKL-QAMNYPYGRKTLFLEIARDIOKAEPEWKILLRYF 186
DB  120 IVFSSSATVYQPEKIPCEEDFKL-QAMNYPYGRKTLFLEIARDIOKAEPEWKILLRYF 179
QY  187 NPVGAHESKGLIGEDPKIPNNLMPYIQQVAVGRITELNLYCHDYPTPTDGSADIRYHMD 246
DB  180 NPVGAHESGLIGEDPKIPNNLMPYISQVAILGKLAQLSVFGSDYDTHDGTGVRDYIHVD 239
QY  247 LADGHIAALRKLFTTENIGCTAYNLGTRGTSVLEMTAFKASGKKIPVKLCPRRPGD 305
DB  240 LAEGHVAAMQ--AKSNVAGTHLNLGSRASSVLEIRAFEAASGLTPIYEVKPRRAGD 296
QY  306 ATEVYASTERAEKELGKANYGVEMCRDWNWAKNPNMGY 346
DB  297 LACFYADPSYAKAIGQVOTQDILQTMEDSWRVVSNPNMGY 337

RESULT 13
GALE HAEIN STANDARD; PRT; 338 AA.
ID  GALE HAEIN STANDARD; PRT; 338 AA.
AC  P24325;
DT  01-MAR-1992 (Rel. 21, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
DE  galactose 4-epimerase).
GN  GALE OR H10351.
OS  Haemophilus influenzae.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC  Pasteurellaceae; Haemophilus.
OX  NCBI_TaxID=727;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=RM 7004 / Serotype B;
RX  MEDLINE=92065797; PubMed=7542800;
RA  Maskell D.J., Szabo M.J., Butler P.D., Williams A.E., Moxon E.R.;
RT  "Molecular analysis of a complex locus from Haemophilus influenzae
RT  involved in phase-variable lipopolysaccharide biosynthesis.";
RL  Mol. Microbiol. 5:1013-1022 (1991).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Rd / KW20 / ATCC 51907;
RX  MEDLINE=95350630; PubMed=7542800;
RA  Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA  Karlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA  McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA  Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA  Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA  Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA  Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA  Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA  Venter J.C.;
RT  "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT  Rd.";
RL  Science 269:496-512 (1995).
CC  -1- FUNCTION: BY CONTROLLING THE INTERNAL GALACTOSE CONCENTRATION,
IT MAY BE LINKED TO THE BIOSYNTHESIS OF LIPOPOLYSACCHARIDE
SURFACE MOLECULES, WHICH ARE IMPORTANT FOR THE PATHOGENESIS OF
H. INFLUENZAE.
CC  -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
CC  -1- COFACTOR: NAD.
CC  -1- PATHWAY: Galactose metabolism; third step.
CC  -1- SUBUNIT: Homodimer (By similarity).
CC  -1- SIMILARITY: Belongs to the sugar epimerase family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  EMBL; X57315; CAA40568.1; -.
CC  EMBL; U32719; AAC22012.1; -.
CC  PIR; A64063; A64063.
CC  HSRP; P09147; IXL.
CC  TIGR; H10351; -.
CC  InterPro; IPR001509; Epimerase_Dh.
CC  InterPro; IPR005886; GALE.
CC  InterPro; IPR000205; NAD_BS.
CC  InterPro; IPR008089; Nuc_sugar_epim.
CC  Pfam; PF01370; Epimerase; 1.
CC  PRINTS; PR01713; NUCEPIMERASE.
CC  TIGRFAMS; TIGR01179; gale; 1.
KW  isomerase; NAD; Galactose metabolism; Complete proteome.
FT  NP BIND 2 33 NAD (POTENTIAL).
FT  VARIANT 285 285 A -> P (IN STRAIN RM 7004).
FT  VARIANT 333 333 P -> S (IN STRAIN RM 7004).
SQ  SEQUENCE 338 AA; 37165 MW; 449934B0E21C8A56 CRC64;

Query Match 49.9%; Score 920.5; DB 1; Length 338;
Best Local Similarity 53.2%; Pred. No. 6.2e-66;
Matches 181; Conservative 56; Mismatches 96; Indels 7; Gaps 3;

QY  8 ILVTGGAGFTGTHVQLLKAGFSVSIIDNFNSVMEAVDRVQVQVGLLSQNLQFTQGD 67
DB  3 ILVTGGAGFTGTHVQLLKAGFSVSIIDNFNSVMEAVDRVQVQVGLLSQNLQFTQGD 58
QY  68 LNRDDLEKLFSTTDFAVIHFAHAGLKAFAESVAKPRYDFNLTGNTINLYEFMAKYNCKK 127
DB  59 ILDRALLQKIPAEINSEINSEINSEINSEINSEINSEINSEINSEINSEINSEINSEINSEIN 118
QY  128 MVFSSSATVYQPEKIPCEEDFKL-QAMNYPYGRKTLFLEIARDIOKAEPEWKILLRYF 186
DB  119 FVFSSSATVYQPEKIPCEEDFKL-QAMNYPYGRKTLFLEIARDIOKAEPEWKILLRYF 178
QY  187 NPVGAHESKGLIGEDPKIPNNLMPYIQQVAVGRITELNLYCHDYPTPTDGSADIRYHMD 246
DB  179 NPVGAHESGLIGEDPKIPNNLMPYISQVAILGKLAQLSVFGSDYDTHDGTGVRDYIHVD 238
QY  247 LADGHIAALRKLFTTENIGCTAYNLGTRGTSVLEMTAFKASGKKIPVKLCPRRPGD 306
DB  239 LAVGHLXALQR--HENDAGLHYNLGTGHSYVLDVMYAFKAKANNITIAVYKLVERSGDI 296
QY  307 TEVYASTERAEKELGKANYGVEMCRDWNWAKNPNMGY 346
DB  297 ATCYSDFSLAKELGKANYGVEMCRDWNWAKNPNMGY 336

RESULT 14
GALE DROME
ID  GALE DROME STANDARD; PRT; 350 AA.
AC  Q9W0E5;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  Probable UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
DE  galactose 4-epimerase).

```

CG12030.
 GN Drosophila melanogaster (fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananidis P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bock J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houch J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "the genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarini H., Krommiller B., Pacle J.M., Park S., Wan X.H.,
 RA Rubin G.M., Celnik S.E.;
 RT "A Drosophila full-length cDNA resource."
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 CC -1- FUNCTION: Catalyzes two distinct but analogous reactions: the
 CC epimerization of UDP-glucose to UDP-galactose and the
 CC epimerization of UDP-N-acetylglucosamine to UDP-N-
 CC acetylglucosamine (by similarity).
 CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -1- COFACTOR: NAD (by similarity).
 CC -1- PATHWAY: Galactose metabolism; third step.
 CC -1- SIMILARITY: Belongs to the sugar epimerase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.

CC -----
 DR EMBL: AE003469; AAF47398.1; -
 DR EMBL: AY058582; AAL1381.1; -
 DR HSSP: F09147; 1XEL.
 DR FlyBase: FBgn0035147; CG12030.
 DR InterPro: IPR001509; Epimerase_Dh.
 DR InterPro: IPR005886; Gale.
 DR InterPro: IPR008089; Nuc_sugar_epim.
 DR Pfam: PF01370; Epimerase; 1.
 DR PRINTS: PR01713; NUCEPIMERASE.
 DR TIGRfams: TIGR01179; gale; 1.
 KW Isomerase; NAD; Galactose metabolism.
 FT NP_BIND 5 36 NAD (POTENTIAL).
 SQ SEQUENCE 350 AA; 38697 MW; 0F14EDCC089ADF5 CRC64;
 Query Match 49.8%; Score 918.5; DB 1; Length 350;
 Best Local Similarity 50.7%; Pred. No. 9.3e-66;
 Matches 177; Conservative 57; Mismatches 104; Indels 11; Gaps 3;
 QY 8 ILVTGGAGFIQTHVTVQLLKAGPSVSIIDNPDNS-----VNEAVDRVRQVVGPLLSQNL 61
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 61
 6 VLVTGGAGYIGSHTVLEMLNAGYVNCVNDLNCAYSSGAKLPEALSRYQEIIG---KKV 61
 QY 62 OPTQGLNRRDLEKLFSTTFDAVIHPAGLKAVAESVAKPRRYFDNLTGVTINLYEFMA 121
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 121
 62 NFYRVDIIDRQVRSVFQEHKIDMVAHPAALKAVGESCRIPLOYHNNMTGTNLLLEMA 121
 QY 122 KYNCKMVFSSSATYVGOEPIKCEEDPKL-QANMPYGRTKLFLEEIARDIOKAEPEWKI 180
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 180
 122 DNNVFKFYSSSATYVGPKEPLPVTBETGCTGCTSPYKTKYFTTEILKDLCKSDKRWAV 181
 QY 181 ILLRYFNPVGAHESKLGEDPKGIPNNLMFYIQOVAVGRLTNLYVGHYPTDGSALRD 240
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 240
 182 VSLRYFNPVGAHISGRIGEDPNENPNLMPIAQAQVAVRRPSLSVIGSDFTHTDGTGVRD 241
 QY 241 YHVMDLADGHIAALRKLFTEENIGCTAYNLTGRTGTSVLEMTAFKASGKKIPVKLC 300
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 300
 242 YIHVLDLAEGHVKALDKLRNLTAEFGFFAYNLGTGVYSLVDWAKAFKASGKKVNYTLVD 301
 QY 301 RPPGDATVYASTAEAEKELCKANYGVEMCRDQWNAKKNPMGYACK 349
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 349
 302 RRSQGVATCYADATLADKKLQWKAERGIDKMCEDTWRMQSQNPNGYANK 350
 RESULT 15
 GALE_NEIMB STANDARD; PRT; 339 AA.
 ID GALE_NEIMB STANDARD; PRT; 339 AA.
 AC P56365; Q59617; Q59624;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 DE galactose 4-epimerase).
 DE GALE OR NMB0064.
 GN Neisseria meningitidis (serogroup B).
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=95020536; PubMed=7934827;
 RA Jennings M.P., van der Ley P., Wilks K.E., Maskell D.J., Poolman J.T.,
 RA Moxon E.R.;
 RT "Cloning and molecular analysis of the gale gene of Neisseria
 RT meningitidis and its role in lipopolysaccharide biosynthesis."
 RL Mol. Microbiol. 10:361-369(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,

RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scariato V., Maignani V., Fizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE IN THE INCORPORATION OF
CC GALACTOSE INTO MENINGOCOCCAL LIPOPOLYSACCHARIDE SURFACE MOLECULES,
CC WHICH ARE IMPORTANT FOR PATHOGENESIS.
CC -!- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: Galactose metabolism; third step.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: Belongs to the sugar epimerase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L20495; AA6535.1; -.
DR EMBL; AE002366; AAF40532.1; -.
DR PIR; S39638; S39638.
DR HSP; P09147; 1XEL.
DR TIGR; NMB0064; -.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; GalE.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01179; GalE; 1.
KW Isomerase; NAD; Galactose metabolism; Complete proteome.
FT NP BIND 3 34 NAD (POTENTIAL).
SQ SEQUENCE 339 AA; 37062 MW; 7D702C44F07DA99B CRC64;

Query Match 49.7%; Score 916.5; DB 1; Length 339;
Best Local Similarity 52.2%; Pred. No. 1.3e-65;
Matches 178; Conservative 56; Mismatches 98; Indels 9; Gaps 4;

QY 8 ILVTGAGFIGHTVYVQLKAGFSVSIIDNPNVMEAVDRVQVVGPELLSQNLOFTQGD 67
DB 4 ILVTGTFIGSHVTVSLKSGHQVVLNLCNSSINILPRLKTITG---QEIPFYQGD 59
QY 68 LNRDLEKLFKSTTFDAVIHFAGLKAVASVAKPRRFDNVLGVTINLYEFMAKYNCK 127
DB 60 IRDREILRFIAENRIDSVIHFAGLKAVGESVAEPKMYDNNVSGSLVLAEMAFAGVFS 119
QY 128 MVFSSSATVYGOPEKIPCEEDFKL-QAMNPYGRTKLFLFEIARDIOKAEPEWKILLRYF 186
DB 120 IVFSSSATVYGPCKVPYTEDMPGDTSPYASKSMVERILTDIQADPRMSMLLYRF 179
QY 187 NPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLFELNLYGHDPTRGSAIRDIHVMD 246
DB 180 NPVGAHESGLIGBQNGIENLLPVCVQVAGKLPQLAVFGDDYPTPDGTGMRDIHVMD 239
QY 247 LADGHTAALRKLFPTTNI-GCTAYNLGCTGTGVLEMTAFEFKASKKIPVKLCRRPGD 305
DB 240 LAEGHVAANQ---AKSNVAGTHLLNLSGRASVLEIIFAFEAASGLTIPYEVKPRAGD 296
QY 306 ATEVYASTERAEKELGWKANYGVEMCRDQWNKNNPWGY 346
DB 297 LACFYADPSVTKAIGWQTQDRDLTQWEDSWRWVSNPNGY 337

Search completed: March 22, 2004, 21:54:12
Job time : 20 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 21:40:15 ; Search time 67 Seconds
(without alignments)
1648.229 Million cell updates/sec

Title: US-09-913-064A-14

Perfect score: 1845

Sequence: 1 MVSSSHLVTGAGFPGTH.....MCRDQWNAKNNPWGAGKP 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvrius.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1530	82.9	351	10 Q8LDN8	Q8LDN8 arabidopsis
2	1522	82.5	353	10 Q9CAM5	Q9CAM5 arabidopsis
3	1328	72.0	360	10 Q8HOB7	Q8HOB7 oryza sativ
4	1229	66.6	348	10 Q8LEA9	Q8LEA9 arabidopsis
5	1229	66.6	351	10 Q8VZ26	Q8VZ26 arabidopsis
6	1222	66.2	348	10 Q9CTW7	Q9CTW7 arabidopsis
7	1203	65.2	350	10 Q8LFW1	Q8LFW1 arabidopsis
8	1196.5	64.9	355	10 Q7XZQ2	Q7XZQ2 zea mays (m
9	1180	64.0	354	10 Q8LNZ3	Q8LNZ3 oryza sativ
10	1159.5	62.8	408	10 Q8H931	Q8H931 oryza sativ
11	1157.5	62.7	447	10 Q9SGX0	Q9SGX0 arabidopsis
12	1046	56.7	338	10 Q81414	Q81414 bacillus ce
13	1023	55.4	338	16 Q81JK4	Q81JK4 bacillus an
14	999.5	54.2	348	13 Q7Z2M6	Q7Z2M6 xenopus lae
15	998	54.1	337	16 Q8EGE0	Q8EGE0 shewanella
16	983	53.3	338	16 Q9KLH0	Q9KLH0 vibrio chol

17	980.5	53.1	339	16 Q87M58	Q87M58 vibrio para
18	977.5	53.0	335	16 Q8DBP1	Q8DBP1 vibrio vuln
19	973	52.7	339	2 Q83W11	Q83W11 streptococc
20	972	52.7	336	16 Q8DCQ4	Q8DCQ4 vibrio vuln
21	971.5	52.7	335	16 Q8VH83	Q8VH83 brucella me
22	970.5	52.6	335	2 Q54385	Q54385 brucella ab
23	970.5	52.6	335	16 Q8GGM3	Q8GGM3 brucella su
24	968.5	52.5	344	5 Q86153	Q86153 dictyosteli
25	966	52.4	339	16 Q97PK2	Q97PK2 streptococc
26	964	52.2	339	16 Q8DNY6	Q8DNY6 streptococc
27	962	52.1	347	5 Q62107	Q62107 caenorhabdi
28	961	52.1	338	16 Q87HS7	Q87HS7 vibrio para
29	951	51.5	364	2 Q5F8B2	Q5F8B2 moraxella c
30	950.5	51.5	335	16 Q82WJ5	Q82WJ5 nitrosomona
31	944	51.2	340	2 Q69132	Q69132 burkholderi
32	943	51.1	364	2 Q9KHV0	Q9KHV0 moraxella c
33	941	51.0	336	16 Q97P16	Q97P16 streptococc
34	927.5	50.3	339	2 Q9X3S6	Q9X3S6 neisseria m
35	927	50.2	340	16 Q8G3V9	Q8G3V9 bifidobacte
36	921.5	49.9	348	16 Q7V4P1	Q7V4P1 prochloroco
37	916.5	49.7	339	2 Q51148	Q51148 neisseria m
38	914	49.5	338	16 Q8XN29	Q8XN29 clostridium
39	913.5	49.5	338	2 Q937X4	Q937X4 edwardsiell
40	911	49.4	341	16 Q8XQW7	Q8XQW7 ralstonia s
41	908	49.2	326	16 Q8DNL9	Q8DNL9 streptococc
42	894.5	48.5	336	2 Q60109	Q60109 versinia en
43	894.5	48.5	338	2 P95527	P95527 pasteurella
44	894.5	48.5	338	16 Q8X942	Q8X942 escherichia
45	894.5	48.5	338	16 Q8FJR9	Q8FJR9 escherichia

ALIGNMENTS

RESULT 1

ID	Q8LDN8	PRELIMINARY;	PRT;	351 AA.
AC	Q8LDN8;			
DT	01-OCT-2002 (TRENBLrel. 22, Created)			
DT	01-OCT-2002 (TRENBLrel. 22, Last sequence update)			
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)			
DE	Uridine diphosphate glucose epimerase, putative.			
GN	AtUG63180/F16M19.8.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosidII; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
[1]	SEQUENCE FROM N.A.			
RP	Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,			
RA	Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;			
RT	"Full-length messenger RNA sequences greatly improve genome			
RT	annotation.";			
RL	Genome Biol. 0:0-0(2002).			
[2]	SEQUENCE FROM N.A.			
RP	Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,			
RA	Feldmann K.;			
RT	"Full-length cDNA from Arabidopsis thaliana.";			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
[3]	SEQUENCE FROM N.A.			
RP	STRAIN=cv. Columbia;			
RA	Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,			
RA	Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,			
RA	Havashizaki Y., Shinozaki K.;			
RT	"Arabidopsis thaliana full-length cDNA.";			
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY085887; AAC63099.1;			
DR	EMBL; AK117913; BAC42551.1;			
DR	GO; GO:0003824; P:catalytic activity; IEA.			
DR	GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.			


```

DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; GalE.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR008089; Nuc_sugar_epim.
DR Pfam; PF01370; Epimerase; 1.
DR PRINTS; PR01713; NUCEPIMERASE.
DR TIGRFAMs; TIGR01179; GalE; 1.
SQ SEQUENCE 351 AA; 38910 MW; 314BC1FAB8DD091EF CRC64;

Query Match 82.9%; Score 1530; DB 10; Length 351;
Best Local Similarity 81.9%; Pred. No. 5.3e-111;
Matches 285; Conservative 24; Mismatches 39; Indels 0; Gaps 0;

QY 3 SSSQHLVTGGAGFGTHTVQLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQLQ 62
DB 4 SVEQNLVTGGAGFGTHTVQLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQLQ 63
QY 63 FTQGLNRDDLEKLFSTKTTTDAVTHFAGLKAFAESVAKPRYPDFNVLGTINLYEFNAK 122
DB 64 FNLGLNRKGDIEKLFSTKTTTDAVTHFAGLKAFAESVAKPRYPDFNVLGTINLYEFNAK 123
QY 123 YNCKMVFSSATVYGOPEKIPCEDEKIQANPYGRKTLFLEIARDIQAEPEWKIL 182
DB 124 YNCKMVFSSATVYGOPEKIPCEDEKIQANPYGRKTLFLEIARDIQAEPEWKIL 183
QY 183 LRYFNPVGAHESGKLGEDPKGIPNNLMPYIQQAVAGRLTELNVYGHDPYTRDGSAIRDYI 242
DB 184 LRYFNPVGAHESGKLGEDPKGIPNNLMPYIQQAVAGRLTELNVYGHDPYTRDGSAIRDYI 243
QY 243 HWMDLADGHIAALRKLFTEENIGCTAYNLGTGRGTSVLEMTATPEKASGKIPVKLCRR 302
DB 244 HWMDLADGHIAALRKLFTEENIGCTAYNLGTGRGTSVLEMTATPEKASGKIPVKLCRR 303
QY 303 PGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNPVGAGKP 350
DB 304 AGDATAVYASTQAEKELGKWKANYGVEMCRDQWNAKNPVGAGKP 351

RESULT 2
Q9CAM5 PRELIMINARY; PRT; 353 AA.
AC Q9CAM5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Uridine diphosphate glucose epimerase, putative, 80611-78786.
GN F16M19.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.I.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koc H.L., Kreneskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maity B., Marziali A.,
RA Millescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,

```

```

RA Utterback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820 (2000).
DR EMBL; AC010795; AAG51599.1; -.
DR FAP; D96657; D96657.
DR HSP; P09147; IXEL.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; GalE.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF01370; Epimerase; 1.
DR PRINTS; PR01713; NUCEPIMERASE.
DR TIGRFAMs; TIGR01179; GalE; 1.
SQ SEQUENCE 353 AA; 39180 MW; 398900C063C2C337 CRC64;

Query Match 82.5%; Score 1522; DB 10; Length 353;
Best Local Similarity 81.4%; Pred. No. 2.2e-110;
Matches 285; Conservative 25; Mismatches 38; Indels 2; Gaps 1;

QY 3 SSSQHLVTGGAGFGTHTVQLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQLQ 62
DB 4 SVEQNLVTGGAGFGTHTVQLLKAGFSVSIIDNFDNSVMEAVDRVQVVGPLLSQLQ 63
QY 63 FTQGLNRDDLEKLFSTKTTTDAVTHFAGLKAFAESVAKPRYPDFNVLGTINLYEFNAK 120
DB 64 FNLGLNRKGDIEKLFSTKTTTDAVTHFAGLKAFAESVAKPRYPDFNVLGTINLYEFNAK 123
QY 121 AKYCKMVFSSATVYGOPEKIPCEDEKIQANPYGRKTLFLEIARDIQAEPEWKI 180
DB 124 AKYCKMVFSSATVYGOPEKIPCEDEKIQANPYGRKTLFLEIARDIQAEPEWKI 183
QY 181 ILLRYFNPVGAHESGKLGEDPKGIPNNLMPYIQQAVAGRLTELNVYGHDPYTRDGSAIRD 240
DB 184 ILLRYFNPVGAHESGKLGEDPKGIPNNLMPYIQQAVAGRLTELNVYGHDPYTRDGSAIRD 243
QY 241 YHVMDLADGHIAALRKLFTEENIGCTAYNLGTGRGTSVLEMTATPEKASGKIPVKLCR 300
DB 244 YHVMDLADGHIAALRKLFTEENIGCTAYNLGTGRGTSVLEMTATPEKASGKIPVKLCR 303
QY 301 RRPDATEVYASTERAELGKWKANYGVEMCRDQWNAKNPVGAGKP 350
DB 304 RRGDATEVYASTQAEKELGKWKANYGVEMCRDQWNAKNPVGAGKP 353

RESULT 3
Q8H057 PRELIMINARY; PRT; 360 AA.
AC Q8H057;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative UDP-glucose 4-epimerase.
GN OSUGB-3.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
[1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare; TISSUE=Immature seed;
RA Suzuki K., Kitamura S.;
RT "Cloning of UDP-glucose 4-epimerase genes in Oryza sativa.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AB097460; BAC41499.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.

```



```
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; Gale.
DR Pfam; PF01370; Epimerase; 1.
DR PRINTS; PRO1713; NUCPIMERASE.
DR TIGRFAMs; TIGR01179; gale; 1.
DR SEQUENCE 360 AA; 39199 MW; 212BC4D210EA4724 CRC64;

Query Match 72.0%; Score 1328; DB 10; Length 360;
Best Local Similarity 70.0%; Pred. No. 3e-95;
Matches 238; Conservative 50; Mismatches 52; Indels 0; Gaps 0;

Qy 8 ILVTGGAGTGHVTVVQLLKAGFSVSIIDNFNSVMEAVDRVQVGVPLLSQNLQFTQGD 67
Db 15 VLVTTGAGTGHVTVVQLLKAGFSVSIIDNFNSVMEAVDRVQVGVPLLSQNLQFTQGD 74

Qy 68 LRNRDLEKLFKSTTFDAVHFAGLKAVASVAKPRRYDFNLTGTLINLYEFMAKYNCK 127
Db 75 LKSKDDMEKVFAXRDAVHFAGLKAVGESVAHQVYNNVAGTWNLYSAMTKYCK 134

Qy 128 MYFSSATVYGOPEKIPCEDEDFKLQAMPYGRTKLFLEIARDIQKAPENKIL 187
Db 135 IVFSSATAYGOPEKIPCEDEDFKLQAMPYGRTKLFLEIARDIQKAPENKIL 194

Qy 188 PVGAHESGKLGEDPKGIPNNLMFYIQQAVAGRLTELVNMGHDPYTRDGSALRDYIHVMDL 247
Db 195 PIGAHSRSGDIGEDPRGIPNNLMFYIQQAVAGRLTELVNMGHDPYTRDGSALRDYIHVMDL 254

Qy 248 AGCHIAALRKLTFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRPGDAT 307
Db 255 ADCHIAALRKLTFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRPGDAT 314

Qy 308 EYVASTERAELKELGWKANYGVEMCRDOWNAKNNPWGYA 347
Db 315 EYVASTERAELKELGWKANYGVEMCRDOWNAKNNPWGYA 354

RESULT 4
Q8LEA9 PRELIMINARY; PRT; 348 AA.
AC Q8LEA9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE UDP-galactose 4-epimerase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation."; 0:0-0(2002).
RL Genome Biol. 0:0-0(2002).
RN SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AT085528; AA62752.1; -
DR GO; GO:0003924; F:catalytic activity; IEA.
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.

DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; Gale.
DR Pfam; PF01370; Epimerase; 1.
DR PRINTS; PRO1713; NUCPIMERASE.
DR TIGRFAMs; TIGR01179; gale; 1.
DR SEQUENCE 348 AA; 38312 MW; AFFC6CB99819316 CRC64;

Query Match 66.6%; Score 1229; DB 10; Length 348;
Best Local Similarity 66.7%; Pred. No. 1.5e-87;
Matches 228; Conservative 50; Mismatches 62; Indels 2; Gaps 2;

Qy 5 SQHILVTGGAGTGHVTVVQLLKAGFSVSIIDNFNSVMEAVDRVQVGVPLLSQNLQFT 64
Db 2 ARNVLVGGAGTGHVTVVQLLKAGFSVSIIDNFNSVMEAVDRVQVGVPLLSQNLQFT 60

Qy 65 QGDLRNRDLEKLFKSTTFDAVHFAGLKAVASVAKPRRYDFNLTGTLINLYEFMAKYN 124
Db 61 QVDLRDRSALEKIFSETKFDVHFAGLKAVGSSVEKPLLYNNNLVGTITLLEVMQAQHG 120

Qy 125 CKKMFSSSSATVYGOPEKIPCEDEDFKLQAMPYGRTKLFLEIARDIQKAPENKIL 184
Db 121 CKNLVSSSSATVYGOPEKIPCEDEDFKLQAMPYGRTKLFLEIARDIQKAPENKIL 180

Qy 185 YFNPVGAHESGKLGEDPKGIPNNLMFYIQQAVAGRLTELVNMGHDPYTRDGSALRDYIHV 244
Db 181 YFNPVGAHESGKLGEDPKGIPNNLMFYIQQAVAGRLTELVNMGHDPYTRDGSALRDYIHV 240

Qy 245 MDIADGHIARLKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRPG 304
Db 241 IDIADGHIARLKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRPG 304

Qy 305 DATEVYASTERAELKELGWKANYGVEMCRDOWNAKNNPWGY 346
Db 300 DAEVYASTERAELKELGWKANYGVEMCRDOWNAKNNPWGY 341

RESULT 5
Q8VZ26 PRELIMINARY; PRT; 351 AA.
AC Q8VZ26;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative UDP-galactose 4-epimerase.
GN AT4G10960.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RA Yanada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene At4g10960 (GI:15236988).";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN SEQUENCE FROM N.A.
RA Yanada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
```



```
QY 185 YFNPVGAHESKLGEDPKGIPNNLMPYIQOVAVGRLTNLVNYGHDPYTRDGS----- 237
DB 244 YFNPVGAHESGRIGEDPCGTPNNLMPYVQVVGRLFNKLYGDTYTKDGTGVRLSHPY 303
QY 238 -----IRDYTHVMDLADGHIATLRLKLTFTTENIGCTA 268
DB 304 TTIRNLNCLNPCIITLTCTCLSLFHMNVHVDYTHVVDLADGHICALQKLDITE-IGCEV 362
QY 269 YNLGTGRGTSVLEMTAFKASGKKIPVKLCPRPGDTEVIYASTERAELGKWKANYG 328
DB 363 YNLGTGRTGTTVLEMTAFKASGKKIPVKYGRPGDAETVYASTERAELNWKANFGI 422
QY 329 EEMCRDOWNAKNPNWGVAGKP 350
DB 423 EEMCRDOWNASNNPFGGSSP 444

RESULT 12
Q81414
ID Q81414 PRELIMINARY; PRT; 338 AA.
AC Q81414;
DT 01-JUN-2003 (TremBrel. 24, Created)
DT 01-JUN-2003 (TremBrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBrel. 25, Last annotation update)
DE UDP-glucose 4-epimerase (EC 5.1.3.2).
GN BC5448.
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=1271630;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelson B.,
RA Kaparat V., Bhaktacharya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Coleman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haselkorn R., Ponstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis."
RL Nature 423:87-91 (2003).
DR EMBL; AE017015; AAP12309.1; -.
DR GO; GO:0016853; P:isomerase activity; IEA.
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; Gale.
DR InterPro; IPR002085; NAD_BS.
DR InterPro; IPR008089; Nuc_sugar_epim.
DR Pfam; PF01370; Epimerase; 1.
DR PRINTS; PR01713; NUCEPIMERASE.
DR TIGRFAMs; TIGR01179; gale; 1.
KW Isomerase; Complete proteome.
SQ SEQUENCE 338 AA; 37270 MW; 93225A3A6ACA3DD4 CRC64;
```

```
Query Match 56.7%; Score 1046; DB 16; Length 338;
Best Local Similarity 56.6%; Pred. No. 2.6e-73;
Matches 192; Conservative 58; Mismatches 83; Indels 6; Gaps 2;

QY 8 ILVTGGAGFTGTHVQLLKAGSVSIIDNFDNSVMEAVDRVQVGPLLSQNLQFTQGD 67
DB 3 ILVTGGAGYIGSHTCVELLNNYKIIIVDNLSSVSIESINRVKBITG----KQFXYKED 58
QY 68 LRNRDDEKLFSTTTDAVTHFAGLKAVASVAKPRRYDFNLVGTINLYEFWAKYCNCK 127
DB 59 VLNRREALDAIFENAEIVAHFAGFVAGSVIAPLTYHHNITSVLVCEVWQXHVCK 118
QY 128 WVFSSATVYQGEKIPCEEDFKLQANPNYGRTKLFLEIARDIQKAEPEWKIILLRYFN 187
DB 119 IFSSSATVYGIPTETPTEEPFLSNPNYGTQKLMIEQIMRDVAFADAENSIALLRYFN 178
```

```
QY 188 PYGAHESKLGEDPKGIPNNLMPYIQOVAVGRLTNLVNYGHDPYTRDGSALRDYTHVMDL 247
DB 179 PFGAHSRIGEDPCGTPNNLMPYVQVVGRLFNKLYGDTYTKDGTGVRDYTHVMDL 238
QY 248 ADGHIAALRLKLTFTTENIGCTAVNLGTGRGTSVLEMTAFKASGKKIPVKLCPRPGDGT 307
DB 239 ANGHVAKLEKVL--RTGVDAYNLGTGMVSVLEMTAFKASGKKIPVKYTERPPGDA 296
QY 308 EYVASTERAEKELGKWKANYVEEMCRDOWNAKNPNWGY 346
DB 297 VCFADASKAKRELGWATRGLEEMCADSWKQSNKNGY 335

RESULT 13
Q81JK4
ID Q81JK4 PRELIMINARY; PRT; 338 AA.
AC Q81JK4;
DT 01-JUN-2003 (TremBrel. 24, Created)
DT 01-JUN-2003 (TremBrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBrel. 25, Last annotation update)
DE UDP-glucose 4-epimerase.
GN GALB-2 OR BA5700.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone G., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahmoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman J.F.,
RA Hazen A., Cline R., Redmond C., Thwaiter J.E., White O., Salzberg S.L.,
RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria."
RL Nature 423:81-86 (2003).
DR EMBL; AE017041; AAP29332.1; -.
DR TIGR; BA5700; -.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; Gale.
DR InterPro; IPR008089; Nuc_sugar_epim.
DR Pfam; PF01370; Epimerase; 1.
DR PRINTS; PR01713; NUCEPIMERASE.
DR TIGRFAMs; TIGR01179; gale; 1.
KW Complete proteome.
SQ SEQUENCE 338 AA; 37845 MW; F24E14DB2B62C2F9 CRC64;
```

```
Query Match 55.4%; Score 1023; DB 16; Length 338;
Best Local Similarity 54.3%; Pred. No. 1.6e-71;
Matches 184; Conservative 63; Mismatches 86; Indels 6; Gaps 2;

QY 8 ILVTGGAGFTGTHVQLLKAGSVSIIDNFDNSVMEAVDRVQVGPLLSQNLQFTQGD 67
DB 3 ILVTGGAGYIGSHTCVELLNNYKIIIVDNLSSVSIESINRVKBITG----KQFXYKEN 58
QY 68 LRNRDDEKLFSTTTDAVTHFAGLKAVASVAKPRRYDFNLVGTINLYEFWAKYCNCK 127
DB 59 VLNRKNNEIFLENNIEAVIHFAFVAGSVETTTPLAYYNNIISAVLCVDMQXHVCK 118
QY 128 WVFSSATVYQGEKIPCEEDFKLQANPNYGRTKLFLEIARDIQKAEPEWKIILLRYFN 187
DB 119 FIFSSATVYGIPTETPTEEPFLSNPNYGTQKLMIEQIMRDVAKADDEWSIALLRYFN 178
```

QY 188 PVGAHESGKLGEDPKGIPNNLMPIYQVAVGRLTETLVNYGHDYPTDRGSAIRDYIHMDL 247
 Db 179 PFGAHSQGRIGEDNGIPNNLMPIYQVAVGRLTETLVNYGHDYPTDRGSAIRDYIHMDL 238
 QY 248 ADGHIAALRLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRRRPDAT 307
 Db 239 AKGHVKALEKVLTK--GIEAYNLGTGKGSVLEMTAFKASGKKIPVKLCPRRRPDAT 296
 QY 308 EVIYASTRAEKELGWKANYGVEMCRDQWNAKNPMGY 346
 Db 297 ICFADVSKAKRELGEAEYGLEENCVDSMRQVNNKNGY 335

RESULT 14
 Q7ZZM6 PRELIMINARY; PRT; 348 AA.
 AC Q7ZZM6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RA initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan A., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX Klein S., Strausberg R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC051601; AAHS1601.1;
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006012; P:galactose metabolism; IEA.
 DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GalE.

DR InterPro; IPR008089; Nuc_sugar_epim.
 DR Pfam; PF01370; Epimerase; 1.
 DR PRINTS; PRO1713; NUCPEIMERASE.
 DR TIGRFAM; TIGR01179; galE; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 348 AA; 38257 MW; 302A809C103AC173 CRC64;
 Query Match 54.2%; Score 999.5; DB 13; Length 348;
 Best Local Similarity 55.9%; Pred. No. 1.2e-69;
 Matches 194; Conservative 50; Mismatches 90; Indels 13; Gaps 4;
 QY 8 ILVTGGAGFIGTHTVQLLKAGFSVSIIDNPDNSVM-----EAVDRVQVVGPLLSQNL 61
 Db 5 VLVTGGGVIQSHCVLELEAGYFVIDNFHNAIRGANNIPESLHRVHDIVG---KTT 60
 QY 62 QFTQGLNRDDLEKLFSTKTTFDVAIHFAGLKAVAESVAKPRRYFDFNLVTINLYEFMA 121
 Db 61 EFEEVDILDRALDKLFKSKHFSVAVLHFAGLKAVGESVQKPLLYKVNLTGTIQLLEVNN 120
 QY 122 KYNCKMVFSSSATVYGOPEKIPCEEDFKL--QANNPYGRTKLFLEETARDIQKAEPEWKI 180
 Db 121 SHGVKNIVFSSSATVYGDNPYLPIDESHFVGGCTNPYKTKYFIEEMIKDCKAEEDWCA 180
 QY 181 ILLRYFPVGAHESGKLGEDPKGIPNNLMPIYQVAVGRLTETLVNYGHDYPTDRGSAIRD 240
 Db 181 ILLRYFPVGAHESGKLGEDPKGIPNNLMPIYQVAVGRLTETLVNYGHDYPTDRGSAIRD 240
 QY 241 YIHVMDLADGHIAALRLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLC 300
 Db 241 YIHVMDLADGHIAALRLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLC 300
 QY 301 RRPDATEVTVASTRAEKELGWKANYGVEMCRDQWNAKNPMGYA 347
 Db 299 RREGDIATCYADPALAKAELGWKAEYGLDRCEDLNRQACNPPTGFS 345

RESULT 15
 Q8EGE0 PRELIMINARY; PRT; 337 AA.
 AC Q8EGE0;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE UDP-glucose 4-epimerase.
 GN GALE OR S01664.
 OS Shewanella oneidensis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
 OC Alteromonadaceae; Shewanella.
 OX NCBI_TaxID=70863;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MR-1;
 RX MEDLINE=22297686; PubMed=12368813;
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
 RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
 RA Meyer T., Trapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
 RA Manuvelan J., Weidman J., Imprim M., Lee K., Berry K., Lee C.,
 RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
 RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
 RT Shewanella oneidensis";
 RL Nat. Biotechnol. 20:1118-1123(2002).
 DR EMBL; AE015612; AAN54719.1;
 DR TIGR; S01664;
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
 DR GO; GO:0006012; P:galactose metabolism; IEA.
 DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GalE.
 DR InterPro; IPR000205; NAD_BS.

